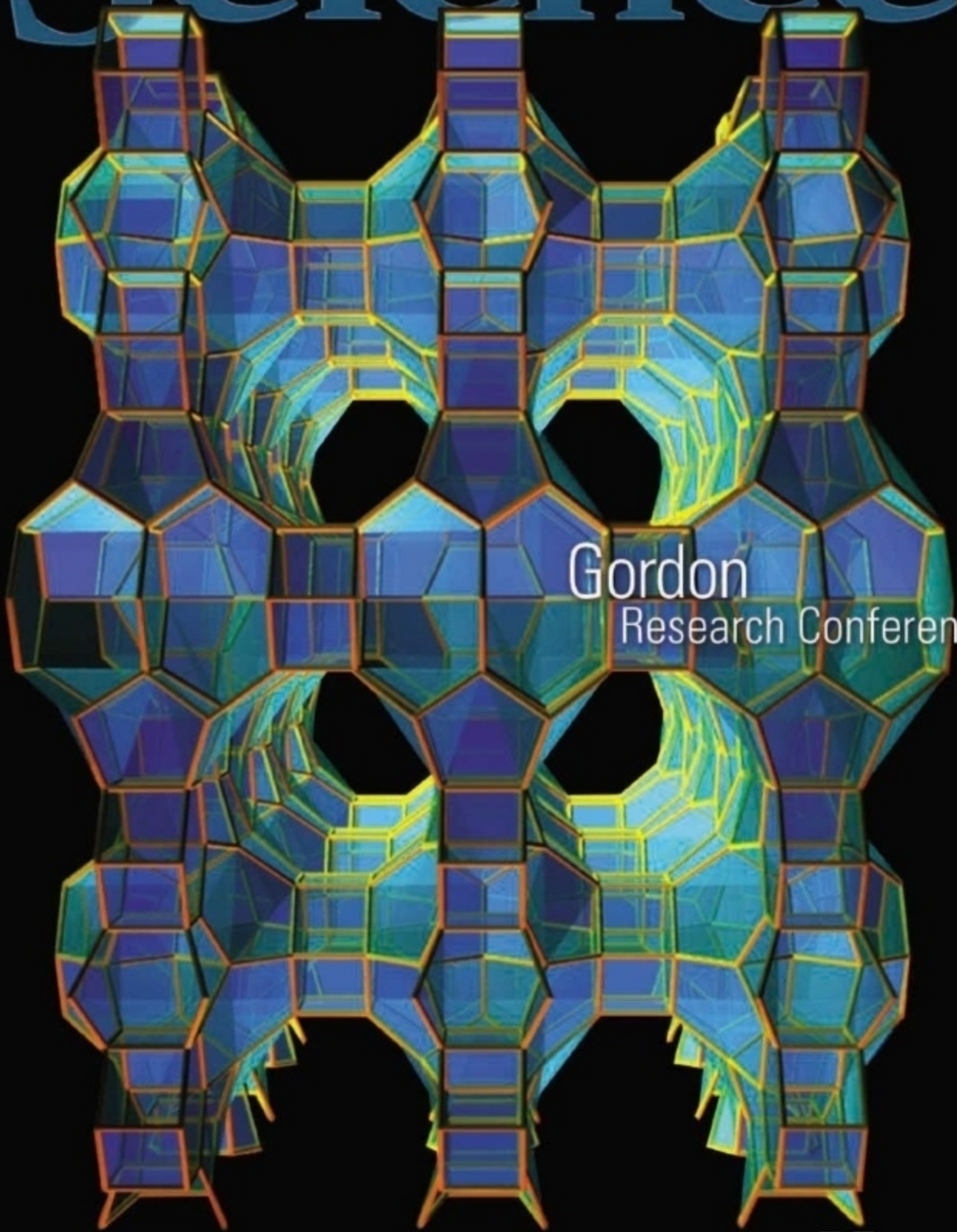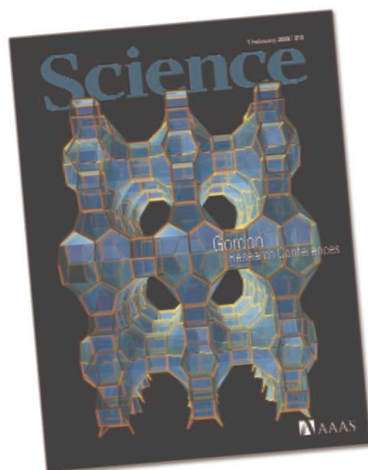# Science

Gordon
Research Conferences

AAAS

COVER

A three-dimensional model of the topological structure of zeolite SSZ-65. The Gordon Research Conference on Nanoporous Materials will be held 15 to 20 June 2008 at Colby College, Waterville, ME. The schedules for the 2008 Gordon Research Conferences begin on page 637.

Model creation and rendering: Kelly Harvey and Scott Harvey

560

579

Science

SCIENCE EXPRESS

www.scienceexpress.org

CLIMATE CHANGE

Human-Induced Changes in the Hydrology of the Western United States
T. P. Barnett et al.
Combining a regional hydrologic and global climate model implies that human-caused CO_2 emissions have already greatly changed river flows and snow pack in the western United States.
10.1126/science.1152538

ASTROPHYSICS

Asphericity in Supernova Explosions from Late-Time Spectroscopy
K. Maeda et al.
Spectroscopic signatures show that supernova explosions of stars that have lost their hydrogen envelopes are strongly aspherical and may be jetlike.
10.1126/science.1149437

GENETICS

High-Resolution Mapping of Crossovers Reveals Extensive Variation in Fine-Scale Recombination Patterns Among Humans
G. Coop, X. Wen, C. Ober, J. K. Pritchard, M. Przeworski
High-density genotyping of individuals from 82 families shows unexpected variation in the number of meiotic crossovers and in the relative activity of recombination hotspots.
10.1126/science.1151851

GENETICS

Sequence Variants in the *RNF212* Gene Associate with Genomewide Recombination Rate
A. Kong et al.
A variant of a human gene associated with high rates of recombination in males and low rates in females is an ortholog of a nematode gene essential for recombination.
10.1126/science.1152422

TECHNICAL COMMENT ABSTRACTS

REVIEW

594

BREVIA

RESEARCH ARTICLE

REPORTS

REPORTS *CONTINUED...*

613

AAAS
ADVANCING SCIENCE, SERVING SOCIETY

Second hand?

SCIENCE**NOW**

www.sciencenow.org DAILY NEWS COVERAGE

Tool Use Is Just a Trick of the Mind
Primate brains learn how to use pliers and other implements
by treating them as part of the body.

Solving the Carbon-14 Mystery
Physicists figure out why the anthropological dating tool
decays so slowly.

The Ocean's Biological Deserts Are Expanding
Global warming may be driving an enlargement of the sea's
least productive regions.

Suppression of
translation by FMRP.

SCIENCE **SIGNALING**

www.stke.org THE SIGNAL TRANSDUCTION KNOWLEDGE ENVIRONMENT

**PERSPECTIVE: Metabotropic Glutamate Receptors and
Fragile X Mental Retardation Protein—Partners in
Translational Regulation at the Synapse**
J. A. Ronesi and K. M. Huber
On the road to protein synthesis–dependent plasticity, FMRP is the
brake and mGluRs are the gas.

EVENTS
Plan to attend a meeting related to cell signaling.

Making connections through collaboration.

SCIENCE **CAREERS**

www.sciencecareers.org CAREER RESOURCES FOR SCIENTISTS

**Maximizing Productivity and Recognition, Part 2:
Collaboration and Networking**
S. Pfirman, P. Balsam, R. E. Bell, J. D. Laird, P. Culligan
Collaboration and networking help make connections that can
advance both science and your career.

What's Ahead for Early-Career Scientists?
B. L. Benderly
A comprehensive examination finds opportunities in the U.S.
brighter in industry than in academia.

Learning to Manage
H. Franzen
A workshop series in Germany teaches management skills
to young scientists before they need them.

February 2008 Funding News
J. Fernández
Learn about the latest in research funding opportunities,
scholarships, fellowships, and internships.

SCIENCE**PODCAST**

Download the 1 February
Science Podcast to hear about
how languages evolve in bursts,
human-induced changes in
U.S. hydrology, the latest on
stem cells, and more.
www.sciencemag.org/about/podcast.dtl

Separate individual or institutional subscriptions to these products may be required for full-text access.

<< Assessing Earth's Inorganic Hydrocarbons

A long-standing question, important not just for petroleum resources but possibly in the origin of life, is the degree that a series of inorganic reactions that lengthen carbon chains (known as Fischer-Tropsch type reactions) might yield hydrocarbons from mantle methane. Although several examples of such hydrocarbons have been inferred, it has been difficult to demonstrate a purely mantle, abiogenic origin in the face of abundant biogenic hydrocarbons. **Proskurowski** *et al.* (p. 604) now show that the abundance of hydrocarbons in the Lost City vent field, an off-axis system in the Atlantic Ocean, decreases systematically with chain length in a manner predicted by Fischer-Tropsch type reactions. Analysis of carbon isotopes further support an inorganic origin. Because this system is likely representative of many similar systems in the oceans, an abundant source of mantle-derived hydrocarbons may be present on Earth, as well as during Earth's early history.

Water as Glass and Liquids

When molecular liquids form glassy phases, the energetic change is often of the same magnitude as when they form crystals—in both cases, large amounts of translational and rotational energy must be lost. In that regard, the glass transition for pure water that occurs at between 120 and 160 kelvin is puzzling in that it occurs with a very modest change in heat capacity. **Angell** (p. 582) reviews the many studies of water's glass transition, including those of aqueous solutions and of water confined to nanoscopic environments. He concludes that ~ 225 kelvin, a temperature often associated with water's "second critical point," an order-disorder transition occurs that accounts for most of the energetic changes. Hence, liquid water appears to exist in two forms—a "fragile" liquid (a poor glass-former) above this temperature, and a "strong" liquid (a good glass-former) below.

More Relaxed on the Surface

The glassy state of materials, in which a liquid-like structure is frozen in place below a specific temperature, may manifest differently in the bulk of the materials versus the surface region. **Fakhraai and Forrest** (p. 600; see the Perspective by **Dutcher and Ediger**) probed the glass transition in an amorphous polymer by embedding gold nanoparticles onto the surface of a polystyrene film and allowing them to sink into the film, where they make small indentations. They then removed the gold, using mercury, and were able to watch the surface relaxation at various temperatures. They observed enhanced surface relaxation, that is, greater mobility of the polymer chains, at the surface relative to the bulk.

Cuprate Liquid Crystal

Recent experimental work has revealed exotic electronically ordered phases in correlated electron systems akin to those seen in conventional liquid crystals. These effects have manifested themselves as anisotropic transport properties, in which conductivity depends upon direction within the sample. **Hinkov** *et al.* (p. 597, published online 10 January) used neutron scattering to investigate the role of spin fluctuations in the macroscopic, nematic liquid-crystalline electronic behavior of the high-temperature superconductor $YBa_2Cu_3O_{6.45}$. They find that an anisotropic ordering of the spins begins at 150 kelvin, well above the temperature where static magnetism occurs, and appears to develop in parallel to the previous reported transport properties. They argue that these fluctuating spins are at the core of the electronic liquid-crystalline behavior in correlated electron systems.
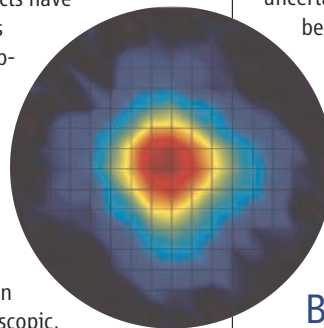
Food for Thought

One of the most potentially harmful effects of climate change may be its impact on agriculture in food-insecure regions. **Lobell** *et al.* (p. 607; see the Perspective by **Brown and Funk**) analyze the climate change–related risk for agriculture in 12 regions worldwide that collectively represent a population of nearly 1 billion people in order to identify which general approaches to adaptation will be most effective in different areas. They find that South Asia and Southern Africa are two regions particularly at risk from negative impacts on several crops, and that uncertainties vary widely by crop. Also, because the reasons underlying a region's vulnerability differ, the adaptation priorities that ultimately need to be followed will depend on how investment institutions perceive uncertainty and risk.

Getting the Balance Right

Balancing selection, the maintenance of multiple alleles within a population, is a means by which genetic diversity may be maintained within a species. **Seidel** *et al.* (p. 589, published online 10 January) have discovered a globally distributed genetic incompatibility that causes embryonic death among natural isolates of the nematode worm, *Caenorhabditis elegans*. The incompatibility persists despite its negative consequences for fitness, which contradicts the prediction that natural selection should eliminate genetic incompatibilities from interbreeding populations.

Regulating Ovulation

In mammals, the ability of a female to remain fertile for an extended period depends on the continuous awakening of primordial follicles from their dormant state in the ovary. Menopause, or the natural end of female reproductive life, occurs when the pool of primordial follicles has been depleted. The mechanisms controlling follicular activation have remained a mystery. **Reddy *et al.*** (p. 611; see the news story by **Marx**) now reveal that follicle activation is controlled by the oocyte PTEN (phosphatase and tensin homolog deleted in chromosome 10)–phosphatidylinositol 3-kinase pathway. In a mouse model where Pten is deleted specifically in oocytes, the entire pool of primordial follicles is prematurely activated and subsequently depleted in early adulthood, which results in premature ovarian failure.

Maternal Influences

Fertilization is a dynamic process of the transition from two highly specialized cells—the oocyte and spermatozoon—into the totipotent zygote. Maternal and paternal contributions to the zygote are not equal. In addition to nuclear DNA, oocytes and spermatozoa are equipped with complementary arsenals of structures such as mitochondria and centrioles for the creation of developmentally competent embryos. By using the microsurgical manipulation of mammalian oocyte nucleolus, **Ogushi *et al.*** (p. 613) demonstrate that the nucleolus, a subnuclear organelle important in ribosome assembly, is exclusively of maternal origin. The oocyte nucleolus is essential for nucleolus assembly in zygotes, and is thus also essential for normal embryonic development.

Growth and Survival, the Complete Toolkit

As tumors progress to a more aggressive state, they acquire multiple genetic alterations, some of which have little functional impact and others that are essential for the continued growth and survival of the tumor cells. **Schlabach *et al.*** (p. 620) and **Silva *et al.*** (p. 617) have developed a functional genomics strategy that will allow, at a genome-wide level, systematic identification of genes required for cell growth and survival. Cell lines derived from human mammary and colorectal cancers and normal mammary tissue showed a similar pattern of so-called "essential" genes, with many residing within functional pathways known to be critical for fundamental cellular processes such as cell cycle and translational control. Importantly, however, additional genes were identified as being essential for the growth of specific cell lines. This functional genomics strategy complements the cancer genome sequencing approaches that have shown recent success and could set the stage for high-throughput discovery of cancer drugs.
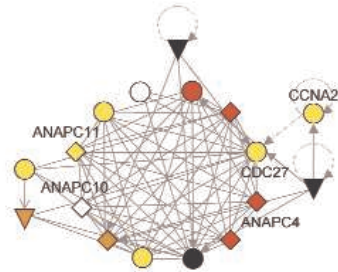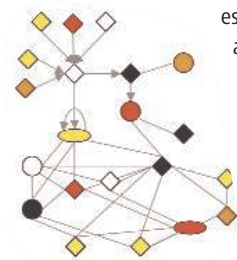
Interfering with Inflammation

The efficient and selective targeting of small interfering RNA (siRNA) molecules to cells could help to harness this technology for treating disease. **Peer *et al.*** (p. 627; see the Perspective by **Szoka**) combine nanoscale liposomal packaging of siRNAs with antibody targeting to immune cells. The targeted siRNA cargo was able to find and efficiently inactivate its target, a key cell-cycle regulating molecule called Cyclin D1. Furthermore, the systemic injection of the packaged siRNA particles reversed pathology in a mouse model of inflammatory bowel disease.

T Cell Role for Cathepsin K

Cathepsins are cysteine proteases that degrade proteins in the lysosome and some cathepsins assist with the processing of antigens for the immune system. **Asagiri *et al.*** (p. 624; see the Perspective by **Krieg and Lipford**) uncover a further but distinct immunological role for another cathepsin, cathepsin K, which is known to be involved in osteoclast function in the bone. Cathepsin K is expressed in immunological dendritic cells and is needed for the complete induction of the inflammatory T helper 17 T cells. In animal models for two autoimmune conditions, pathology was ameliorated by cathepsin K deficiency because of its unexpected involvement in signaling through the innate immune receptor TLR9.

CREDIT: SCHLABACH ET AL.

Donald Kennedy is the Editor-in-Chief of *Science*.

The Real Debate

WE IN THE UNITED STATES ARE SLIDING DOWN A RAMP THAT WILL TAKE US, IN JUST 4 DAYS, to the much anticipated "Super Tuesday" in the presidential nomination cycle, when voters in over 20 states participate in preliminary elections to select their favorite candidate. I have prepared for this by watching, in alternating stages of boredom and disbelief, the numerous "debates" staged by the creative powers who run television. I wonder whether the same sensations haven't affected our scientific colleagues in other nations, where leadership is decided in an atmosphere that is, well, a bit more stately. Here it may be too late to change anyone's mind about their vote on 5 February, but perhaps between now and the culminating summer conventions that will announce the final party candidates, we can have a debate focusing on the candidates' views about science and technology.

I disclaim any intellectual property rights to this idea; probably most of you have already thought of it. My News colleagues at *Science* have already examined the candidates' records and statements (4 January 2008 issue). But a public debate on science could launch disagreements among the candidates and sharpen positions. Chris Mooney and Shawn Otto have organized a group of concerned scientists, journalists, and leaders of government, nongovernment, and business institutions to push for that (www.sciencedebate2008.org). The American Association for the Advancement of Science, the publisher of *Science*, has agreed to cosponsor the debate, and the project has been endorsed by Congressman Bart Gordon (D-TN), chair of the House Science and Technology Committee. In a different but related effort, Research!America invites voters, through a multi-state ad campaign, to contact the candidates and urges the candidates to get out their positions on health and research (www.yourcandidatesyourhealth.org). And Student Pugwash USA and the Federation of American Societies for Experimental Biology are making similar plans.

If we had a science debate among the party candidates, who else might be involved? There are several good science journalists who could moderate (I won't name them because it would make me more enemies than friends). We could pick a scientist as well, but an alternative might be a public figure with a serious interest in science and science policy—someone along the lines of Alan Alda, perhaps? And it would have to be televised. I hope we'd enlist an organization whose style more closely resembles that of the Public Broadcasting Service's NewsHour rather than the YouTube/Cable News Network combination.

Finally, we'd need some questions. In an appearance on National Public Radio's Science Friday (11 January), Shawn Otto urged scientists to submit questions. Here are some of theirs and some of mine:

•What consideration should be given to political affiliation in the appointment of members of advisory committees whose role is to evaluate research quality?

•The president has a Science Adviser who also heads the Office of Science and Technology Policy (OSTP). What attributes would you seek in your Science Adviser, and what kinds of issues would you bring to OSTP?

•What balance would you seek in federal science funding between major-program project research and investigator-initiated basic research grants?

•The budget of the National Institutes of Health was doubled but has decreased for 3 years because its appropriations have been in constant dollars. Would your Administration propose adding inflation costs to that budget in future years?

•If a threatened species exists on private land, does the Endangered Species Act require certain duties of the landowner? What are these, and would you favor changes in the law to alter them?

•In view of public concerns about global warming, are you committed to the mitigation of greenhouse gas emissions? Would you choose a cap-and-trade program or a carbon tax? Why?

•Would you make a commitment to ensure public access to findings made by government scientists in the course of exercising their agency responsibilities?

•Crops derived from recombinant DNA technology are in increasing use in agriculture. Do you favor more intensive regulation to eliminate their possible interference with surrounding natural ecosystems?

In case we can bring this thing off, get your questions ready!

– **Donald Kennedy**

CREDIT: GETTY IMAGES

neurons are central in the emergence from (though not the induction of) an anesthetized state. Both isoflurane and sevoflurane reduced the percentage of active orexin neurons to the levels seen during non–rapid eye movement sleep, yet an orexin receptor antagonist surprisingly did not change the rate of entry into anesthesia. Nevertheless, the same antagonist did markedly delay recovery from an anesthetized state, and a similar delay was observed in orexin-deficient mice and also is seen in some human narcoleptic patients. — GJC

Proc. Natl. Acad. Sci. U.S.A. **105**, 10.1073/pnas.0707146105 (2008).

ECOLOGY/EVOLUTION
The Largest of the Small

The moist forests of the Udzungwa Mountains in south-central Tanzania have yielded an astonishing number of previously undescribed vertebrate species during the past decade. The latest of these, reported by Rovero *et al.*, is a remarkable new elephant shrew or sengi, named *Rhynchocyon udzungwensis*. Related neither to elephants nor to shrews (being much smaller than the former and much larger than the latter), the elephant shrews are an order of mammals that appear to have evolved hardly at all since the Miocene. The new species is the largest sengi of all, weighing in at an average of 700 g and measuring half a meter from the elongated snout to the tip of its tail. On the basis of sighting frequency, Rovero *et al.* estimate a total population of 15,000 to 24,000 occupying an area of 300 km^2. This giant sengi lives in mountain forests 1000 m above sea level; its habitat, along with those of other endemic species of the Udzungwa Mountains, is currently protected and relatively little disturbed by humans. The discovery of yet another new species is a further confirmation of the conservation value of these mountains. — AMS

J. Zool. **274**, 10.1111/j.1469-7998.2007.00363.x (2008).

BIOCHEMISTRY
Assemble Before Use

At sites of vascular injury, a large multimeric glycoprotein (von Willebrand's factor; VWF) secreted from endothelial cells binds platelets to form a hemostatic plug. Within endothelial cells, VWF multimers pack as ordered tubules into cigar-shaped secretory granules called Weibel-Palade bodies; this packaging is essential for orderly secretion of VWF filaments.

Using only the N-terminal propeptide D1D2 and the adjacent D'D3 domains, Huang *et al.* have reconstituted in vitro the formation of Weibel-Palade body–like tubules, which occurs at low pH (6.2) and in the presence of Ca^{2+}. Electron microscopic reconstruction showed that the tubules formed a right-handed helix with 4.2 units per turn, with the repeating unit containing a D'D3 dimer and two propeptides. The authors suggest that these domains form the core of the tubules, with the remaining C-terminal portion of the protein decorating the outside. In the Golgi, the relatively
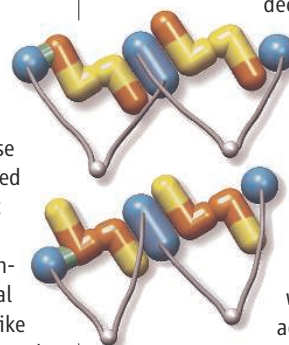
Disulfide-linked (gray) assemblies of D1 (yellow), D2 (orange), and D'D3 (blue).

acidic pH and high Ca^{2+} would increase the interaction between D1D2 and D'D3, juxtaposing the two D3 domains and facilitating intersubunit disulfide bond formation and multimerization. The higher pH in blood would weaken these interactions and allow the helical tubules to unfurl without tangling. — VV

Proc. Natl. Acad. Sci. U.S.A. **105**, 482 (2008).

MATERIALS SCIENCE
The Order of Ordering

When polymers partially crystallize from the melt state, they often pack by formation of lamellae, or stacks of folded, ordered chain segments separated by regions of noncrystalline material. Prior studies suggested that crystallization might be occurring through a spinodal-assisted ordering, associated with the formation of a smectic-like liquid crystalline phase that preceded the formation of the first nuclei. Panine *et al.* used high-brilliance x-ray scattering to look at the earliest stages of crystallization in isotatic polypropylene. They found ordering in the wide-angle x-ray scattering (WAXS) data before the small-angle x-ray scattering (SAXS) data, thus supporting the formation of ordered nuclei at the local scale before more global, liquid crystalline–like ordering. This finding contrasted with earlier SAXS before WAXS data used to support the spinodal hypothesis, which may have been due to detector limitations. More intriguing is that the earliest SAXS peaks support a smectic-like ordering of the nuclei (rather than just the

polymer chains); thus, at these intermediate times, the nuclei may arrange into an ordered structure before growing into the large well-organized lamellae as proposed in a number of recent simulation studies. — MSL

Polymer 10.1016/j.polymer.2007.12.026 (2008).

PHYSIOLOGY
Waking Up to Orexin

What do narcolepsy and anesthesia have in common? Almost a decade ago, a mouse model for human narcolepsy was developed on the basis of results demonstrating that the neuropeptide orexin promoted wakefulness and that genetic ablation of orexinergic neurons yielded mice with behavioral and physiological symptoms remarkably like those of narcoleptic humans. Anesthesia, on the other hand, can be induced by a wide variety of agents such as isoflurane or sevoflurane but has resisted efforts to identify its neural loci of action.

In their mouse model, Kelz *et al.* find that the neural systems innervated by orexinergic

MT. BLANCO FOSSIL MUSEUM PRESENTS
LONE ★
THE WORLDS LARGEST 4-TUSK MASTODON

Mastodon on the Block

A creationist museum in Texas has sold a giant four-tusked skull of a mastodon to help pay its bills.

The skull was found in a gravel pit in 2004. Joe Taylor, founder of the Mt. Blanco Fossil Museum near Lubbock, spent 9 months chipping it out of a block of sandstone. Mastodons went extinct about 10,000 years ago, but Taylor, who doesn't trust radiocarbon dating, believes the fossil is 3000 to 4000 years old.

Dubbing it "Lone Star," he put the skull on display in 2005. But his museum has fallen on hard times, and he put the beast up for sale. At a 20 January auction in Dallas, it went to an anonymous bidder for $191,200, a record sum for a mastodon. All male mastodons had four tusks, but, unlike Lone Star, they often lost their lower ones in fights.

The unusual sale has researchers fuming at another instance of the commercialization of fossils but relieved that the specimen won't be on display anymore. "If you wanted to take the two things that piss off paleontologists the most—the sale of fossils and creationist museums—here we have the both of them," says Thomas Holtz of the University of Maryland, College Park.

Resentment Kills

Bottling up anger can shorten your life, an unusual long-term study of married couples in Michigan concludes.

The study covers 192 couples in the Tecumseh Community Health Study who were between the ages of 30 and 69 in 1971. To classify the men and women as anger "suppressors" or "expressers," researchers asked each of them to imagine getting chewed out by a police officer or spouse for something he or she hadn't done. Anger suppressors were those who failed to protest unfair attacks or felt guilty later if they had gotten mad.

The researchers, led by psychologist Ernest Harburg, now a professor emeritus at the University of Michigan School of Public Health in Ann Arbor, tracked mortality in the couples over 17 years, controlling for age, smoking, weight, blood pressure, education, and heart and lung problems. Among couples in which both members were anger suppressors, the mortality rate was twice that of the other groups combined, the researchers report in the January issue of the *Journal of Family Communication*. Twenty-six of the couples (14% of the sample) were in this category; there were 13 deaths, compared with 41 in the remaining 166 pairs.

Research on the "dyadic relationship" as a unit is rare, says Harburg, who adds that two anger suppressors seem to have a synergistically morbid effect on each other. Psychologist Janice Kiecolt-Glaser of Ohio State University College of Medicine in Columbus says the data "add weight to the growing evidence that poor emotional housecleaning has health consequences in marriages."

If It Ain't Broke …

Scientists have discovered what they say is the oldest fossil of a horseshoe crab. It dates back 445 million years, and it looks very like the ones that ply the North Atlantic today.

The old and the new.

Called *Lunataspis*, the fossil was found during recent excavations in late Ordovician deposits on the coast of Manitoba by scientists from the Manitoba Museum in Winnipeg and the Royal Ontario Museum in Toronto, Canada.

Reporting in the latest issue of the British Palaeontological Association journal *Palaeontology*, David Rudkin of the Royal Ontario Museum and colleagues say the horseshoe crab is a remarkable example of evolutionary stasis. Like cockroaches, they are "the quintessential 'living fossils' of biology text-books."

The new fossil's fused body segments, long tail spine, and large crescent-shaped head shield with compound eyes closely resemble those of today's horseshoe crab, thus pushing back the record of this animal's body plan by more than 100 million years. Until now, the oldest one on record has been from the 320-million-year-old Bear Gulch deposits of Montana.

"This is a unique flag pin in the history and evolution of the horseshoe crabs," says Lyall Anderson, an expert on fossil arthropods at the University of Cambridge's Sedgwick Museum of Earth Sciences. Now, he says, it's up to someone to find the common ancestor of the xiphosurids (as these crabs are called) and the other group of extinct horseshoe crabs called the synziphosurines. That should be somewhere in Cambrian times, more than 500 million years ago.

Cold Comfort

Layovers in Paris just got more interesting. Until 15 March, the city's two main airports are home to a photo exhibit about polar research at the National Centre for Scientific Research—including this shot of a Dolgan hunter and reindeer breeder in Siberia.

<< In the Movies

SHIMMER. The raging sea in *Pirates of the Caribbean* isn't really made of water, but it's still a knockout. Its creators—Stanford University computer scientist Ronald Fedkiw and Frank Petterson and Nick Rasmussen of Industrial Light and Magic—will receive an Oscar next week for advancing the science and technology of special effects.

The liquid terminator in *Terminator 3*, the sea in *Poseidon*, and similar watery effects in nearly 20 other blockbuster films all rely on a fluid-simulation method developed in Fedkiw's lab in the early 2000s as a tool for computational physics, with funding from the U.S. Office of Naval Research. It generates smoother and more detailed fluids than its predecessors.

Petterson and Rasmussen have applied the method, now used by the Navy to simulate explosions, to filmmaking. The method's big splash on the silver screen "was just luck," Fedkiw says.

FACT AND FICTION

QUID PRO QUO. Here's an idea for drawing attention to your research: Lend a hand with promoting a movie whose plot is tangentially linked to what you study. Jeff Kimble of the California Institute of Technology in Pasadena and Max Tegmark and Edward Farhi of the Massachusetts Institute of Technology (MIT) in Cambridge are doing just that by helping to create a buzz about *Jumper*, whose protagonist can "teleport" himself instantly through space by dint of special mental powers.

In 1998, Kimble performed a less dramatic form of teleportation by transferring the quantum state of one photon to another a distance away. He has volunteered to field phone calls from reviewers and journalists about real teleportation research. Theorists Tegmark and Farhi participated in a 16 January panel discussion hosted by an MIT student group at which scenes from the movie were previewed.

All three researchers say they were concerned that their participation could be misinterpreted as endorsing the scientific validity of the movie. "There's benefit and risk," Kimble says, "but I think there is a responsibility to convey to the public what's really going on in science." To maintain independence, Kimble declined compensation. Tegmark and Farhi insisted on a "no holds barred" discussion—but accepted an undisclosed appearance fee.

"I told [the promoter], 'You're promoting a $100 million movie; I want to be paid,'" Farhi says. The movie premieres on 8 February.

IN BRIEF

TAPPING METHANE. Research teams at Cardiff University in the U.K. and Northwestern University in Evanston, Illinois, are the inaugural winners of the Dow Methane Challenge. The chemical manufacturing giant started the competition last year to promote research on converting methane into a raw material for chemical feedstocks, which could open the door to tapping natural gas reserves around the world for chemicals and alternative fuels. The two groups, whose proposals to work on the problem were chosen from nearly 100 submissions, will share $6.4 million to do research in collaboration with Dow's chemists and engineers.

Got a tip for this page? E-mail people@aaas.org

Celebrities

FAME INFLATION. Forest ecologist Steven Running isn't a Nobel laureate. But try telling that to his employer, the University of Montana, Missoula.

Running was feted as a recipient of the 2007 Nobel Peace Prize at an all-university lecture shortly after the awardees were announced in October. Then there was the "2007 Nobel Peace Prize" engraved on his parking space, and, his favorite, his 2007 Nobel Peace Prize bike rack. Eventually, reporters—from the local newsweekly to the *New York Times*—were calling him a Nobel laureate.

Why all the confusion? Running was a leading participant in the fourth assessment of the state of climate science issued by the Intergovernmental Panel on Climate Change (IPCC), which shared the Nobel Peace Prize last fall with Al Gore. But something like a couple of thousand scientists have done much the same IPCC work over the past 20 years. "Trying to explain that to the press doesn't work," he says. "People just don't make the distinction between real Nobel Prize winners and IPCC."

He speculates that his notoriety may be due in part to the fact that, around thinly populated Montana, he's "about the closest thing they ever had" to a Nobel Prize winner. Still, he's found an effective way to deflate people's misperceptions. He just tells them: "No, I don't get any of the money."
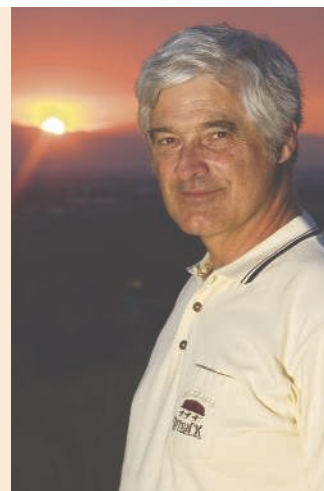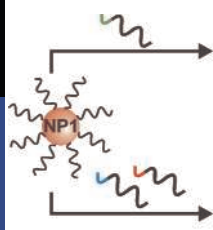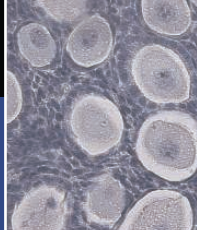
ENERGY RESEARCH

DOE's Disappointing Budget Makes It Harder to Stick to the Basics

Malcolm Stocks has spent his career scrutinizing the arrangement of electrons in alloys, magnetic materials, and nanostructures and developing techniques to calculate electronic structures. Even experimental physicists may find the work a bit esoteric. But this year, Stocks, 64, was looking forward to tackling a real-world problem: improving the efficiency of the next generation of nuclear reactors by predicting more accurately how the crystalline structure of nuclear fuel changes as the material "burns." "I'm really turned on about doing science that impacts real materials and real technological problems," says Stocks, a theoretical physicist at Oak Ridge National Laboratory in Tennessee.

Then Congress intervened. Stocks and his team had applied to the U.S. Department of Energy (DOE) for a $3 million, 3-year grant, and Stocks knew that his odds of success were only about one in three. But they shrank to zero after legislators wiped out most of the requested increase for DOE's Office of Science as part of an omnibus spending bill to fund most government agencies for the rest of the 2008 fiscal year (*Science*, 4 January, p. 18).

Once they received the bad news—a 2% boost rather than the 20% requested, a difference of $220 million—officials within the Basic Energy Sciences (BES) program pulled the plug on an $80 million initiative that had attracted more than 700 proposals, including one from Stocks's team, to pursue the "use-inspired" science of solar energy, hydrogen fuels, and advanced nuclear reactors. (Some 40 grants had been funded last spring in a first round that didn't cover nuclear research.) The grant would have supplied about 25% of Stocks's funding and would have enabled him to expand his 12-member group. "This would have brought in a new member of the staff in a new area," he says.

As the largest of the Office of Science's six divisions, BES funds research in materials sciences, chemistry, condensed matter physics, and related fields. Its $1.28 billion budget also pays for synchrotron x-ray sources, neutron sources, and other "user
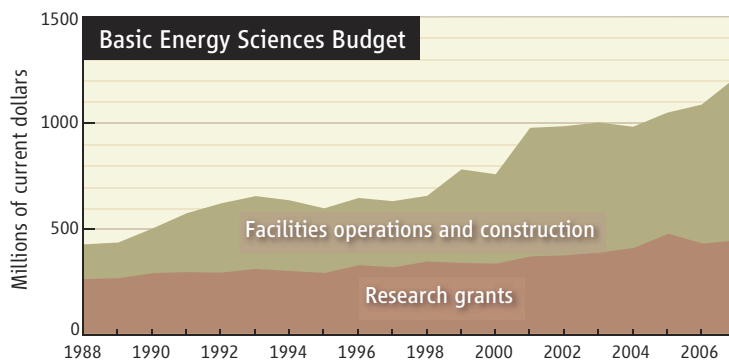
Falling behind. Funding for research has not kept pace with the cost of building and running user facilities.

facilities" at the Office of Science's 10 national labs. The 2008 budget will not lead to major layoffs at those labs, as have cuts to DOE's particle physics budget (*Science*, 11 January, p. 142). But it will mean an array of smaller cost-saving measures that will have an impact on science. For example, user time will be cut by up to 20% at BES's already-oversubscribed user facilities, which support thousands of university researchers. "Nationwide, it is a very significant impact," says Thom Mason, director of Oak Ridge. "But it's not manifest in a dramatic way at one lab."

Wayne Hendrickson, a protein crystallographer at Columbia University and the Howard Hughes Medical Institute in Chevy Chase, Maryland, worries about the long-term implications of BES's suddenly leaner budget. Hendrickson studies macromolecules at both the Advanced Photon Source at Argonne National Laboratory in Illinois and the National Synchrotron Light Source at Brookhaven National Laboratory in Upton, New York, where his group maintains two beamlines. Staff at the labs will strive to keep the machines running as much as possible, Hendrickson says. But he worries that to do that, they will have to cut back on maintenance and personnel. "There are effects that may not be felt immediately."

Some observers are concerned that users will head for overseas facilities, where queues may be shorter. Hendrickson says he would consider going back to Europe, where he ran experiments in the early 1990s. "We don't want to get on an airplane, but if we have to, we will," he says. If things don't improve next year, Hendrickson says, he might do it again to avoid falling behind the competition.

The cuts in the BES budget are also hampering plans to build some user facilities. Researchers at Brookhaven received only $30 million of the $45 million requested to design and procure parts for the $912 million National Synchrotron Light Source II, which is scheduled to start up in 2015. And researchers at Lawrence Berkeley National Laboratory in California received only $5 million of $17.4 million requested to begin construction of a 28,000-square-meter building that would provide lab space to users of the adjoining Advanced Light Source (ALS).

Construction of the ALS User Support Building will be delayed by at least a year, says Joseph Harkins, project director at the Berkeley Lab. And the project will likely exceed its original $32 million budget. Berkeley Lab officials will ask DOE to cover the increased expense, Harkins says. If no extra money is available, then the capabilities of the new building would have to be reduced, Harkins says.

Many researchers fret most about the cuts in grants to individuals and small teams at universities and national labs. BES funds them through its $450 million core research program, which has languished in recent years. In the past 2 decades, the overall BES budget has tripled, but most of that increase has gone to building and running user facilities (see graph, above). The use-inspired grants would have helped restore some ▶

FOCUS

Where next with
stem cells?

560

Space science and
the next president

564

balance, many researchers say.

The use-inspired grants were also designed to link basic research within the Office of Science with applied research in DOE's offices of Energy Efficiency and Renewable Energy, Nuclear Energy, and Fossil Energy. Since 2002, BES has hosted 11 workshops to identify the fundamental questions that must be answered to achieve

major advances in various energy technologies. The proposals came in response to calls issued after workshops on solar energy, hydrogen fuels, and nuclear energy.

Even if the budget is better for the 2009 fiscal year that begins on 1 October, nobody expects DOE to simply pick up where it left off. Applicants had to survive intralab competitions before submitting their proposals to

DOE, and the entire process will likely have to be repeated. What he did last year, Stocks says, "is now an enormous waste of time."

Still, many scientists say that if DOE issues a new call for use-inspired research proposals, they will answer. "If we think it's important science," Stocks says, "then of course we're going to apply."   –ADRIAN CHO

With reporting by Robert F. Service.

NUTRITION SCIENCE

Lancet and MSF Split Over Malnutrition Series

They both have the interests of malnourished children at heart. But in a nasty spat about a series of scientific papers, Médecins sans Frontières (MSF), the international charity, and the medical journal *The Lancet* are accusing each other of damaging that very cause. The flap, which centers on the merits of so-called ready-to-use therapeutic foods, has become so heated that Richard Horton, editor of *The Lancet*, says for now he will no longer accept articles by MSF staffers. Many outsiders are calling on the two to stop bickering and focus on the plight of malnourished children instead.

At issue is a series of five articles about undernutrition produced by a special study group led by five leading nutrition scientists. Funded by the Bill and Melinda Gates Foundation, the group explored the causes and consequences of malnutrition and examined the scientific evidence for various interventions. The series was unveiled at press conferences in seven cities around the world on 16 January. It is expected to have a major policy impact, for instance, at the next meeting of the United Nations System Standing Committee on Nutrition (SCN), in March in Hanoi, Vietnam.

But the package had barely gone online, along with three commentaries, when MSF staffers in Geneva, Switzerland—who had read embargoed copies of the papers—published a harsh critique on their Web site. Their core complaint is that the series devotes little attention to ready-to-use therapeutic foods. Made under names such as Plumpy'nut, these peanut-based products, high in energy and protein, can be used at home; they are widely used against severe acute malnutrition by MSF and other aid organizations, especially in emergency situations. They have

"transformed" practices, says Tido von Schoen-Angerer, head of MSF's Campaign for Access to Essential Medicines, and should be used much more widely.

"By failing to strongly endorse" that strategy, "*The Lancet* authors are undermining the support for this lifesaving intervention," MSF's statement reads. MSF decided to issue

Food fight. MSF contends *The Lancet* ignored the value of the ready-to-use food, like Plumpy'nut, that the group is pushing.

the statement instead of writing a letter to the editor because "we felt it was important to have a response immediately," says Von Schoen-Angerer.

But study group member Zulfiqar Bhutta of Aga Khan University in Karachi, Pakistan, says the criticism is "completely misplaced." Bhutta e-mailed *Science* a statement on behalf of the researchers that points out that one of the papers did note that treating severe malnutrition at home "is now possible and has been recommended." The problem, the statement continues, is that the authors could not identify any randomized controlled clinical trials

investigating the food's effect on mortality.

To Horton, the fact that "MSF has punctured the beginning of an advocacy campaign based on the best science" is "unforgivable." The result, he says, is that the fight, rather than child malnutrition, will get most of the attention. Horton says that *The Lancet* has "put our relationship with MSF on hold [which includes a temporary ban on papers by MSF authors] until I have a clear response about how this could have happened." He says he has received e-mails from key MSF employees apologizing for their organization's behavior; MSF staffers who asked not to be named confirmed to *Science* that the issue has divided the organization. Geoff Prescott, director of MSF in the Netherlands, says, "I thought the language in the statement was a bit strong."

Still, MSF has a point, says SCN Secretary Roger Shrimpton. Rigorous clinical trials in nutrition are often hard to do, especially in the areas where organizations like MSF operate. "If we waited for randomized controlled trials for everything, we'd do only half of what we're doing," he says. However, "why MSF needs to make such a hullabaloo, I'm not quite sure," Shrimpton adds. "This is a fantastic series. It's the beginning of a process; it's not the Bible." *The Lancet* and MSF should mend fences as soon as possible, he says.

–MARTIN ENSERINK

CREDIT (LEFT): JOHN SCHULTS/REUTERS/LANDOV

INTELLECTUAL PROPERTY

Indian Government Hopes Bill Will Stimulate Innovation

The Indian government is preparing to introduce legislation that it hopes will reverse the traditional hands-off attitude at most Indian universities toward commercializing the results of basic research. The proposed bill, a draft of which was obtained by *Science*, sets out rules that institutions must follow once their scientists make a patentable discovery. In addition to serving as the first such nationwide guidelines, the legislation is meant to send a message to university officials that technology transfer is part of their job.

"The idea is to create an environment of innovation," says Maharaj Kishan Bhan, secretary of the Department of Biotechnology, which helped to draft the bill. Until now, he says, "university administrators [have been] free to encourage or discourage patenting and commercialization efforts. And there are many who believe that industry and academia should be kept miles apart."

The legislation directs institutions to report patentable discoveries to the funding agency as soon as they come to light and to file patent applications within a year. Institutions and inventors who fail to meet the deadline would forfeit their intellectual-property rights to the agency that funded the research. Institutions must give inventors at least 30% of any revenues from a

Off the bench. Manju Ray wants to translate her research into therapies for cancer patients.

patent and spend the rest on research.

Those rules might have helped biochemist Manju Ray after she developed an experimental drug based on the anticancer properties of methylglyoxal—a metabolic byproduct found in most organisms, including humans. She works at the Indian Association for the Cultivation of Science (IACS) in Kolkata, which took no steps to patent the compound. Finally, Ray agreed to partner with Dabur India Ltd., an Indian pharma-

ceutical company, making it co-owner of domestic and international patents granted in 2003. But Dabur has since lost interest in the therapy and has been pursuing other candidate drugs. Ray cannot license it to another company without Dabur's approval, and IACS has no financial incentive or authority to get involved.

Somenath Ghosh, head of the National Research Development Corp. in New Delhi, which markets technologies developed through public-funded research, expects the bill to pave the way for more governmental support to labs and universities to help with patent filings. But Ghosh and other analysts also wonder whether the government has the resources to carry out its threat to claim patentable discoveries if institutions drag their feet.

Even without the bill, the culture at institutions may be changing. Last year, Ray's institute set up a technology-transfer office, which helped Ray file a patent application for a new formulation of her cancer drug. If she receives the patent, she will be able to explore partnerships with other pharmaceutical companies. "I want to forget about the past and look to the future," she says.

The government plans to introduce the bill this spring in Parliament, where it is expected to win swift passage.

–YUDHIJIT BHATTACHARJEE

NONPROLIFERATION

Dutch Revise Policy Blocking Iranian Students

The Dutch government this week backed away from an antiterrorism policy that had led one university to reject applications from Iranian students and triggered a loud protest among academics. But researchers complain that the revised policy will still make it hard for Iranian scholars and students to study in the Netherlands, and they fear that such policies could spread throughout Europe.

The original policy was the government's attempt to implement a 2006 United Nations resolution that asks all nations to "prevent specialized teaching or training of Iranian nationals … [in] disciplines which would contribute to Iran's proliferation [of] sensitive nuclear activities and development of nuclear weapon delivery systems." Last fall, the Dutch education and foreign affairs ministers told all universities to

exercise "vigilance" in admitting Iranian students. In December, the University of Twente in Enschede announced that it would no longer accept Iranian students because the Dutch Immigration and Naturalization Service (INS) had asked for a guarantee that Iranians on campus would not gain any sensitive knowledge. Officials at the Eindhoven University of Technology said they would consult Dutch intelligence officials while considering Iranian applicants for admission.

Academics and students protested the new policy, calling it overly broad and discriminatory. Their objections were heard: This week, Twente officials said that INS has agreed to withdraw its demand for a guarantee and that the university would reopen its doors to Iranians. Robert Dekker, a foreign ministry spokesperson,

says the government still intends to implement the U.N. resolution by barring Iranian students from admission to certain fields. (Students already enrolled face no such restrictions.) "The ministries and the universities are discussing which studies might fall under the resolution," Dekker told *Science*. The exclusion could include degree programs that are not directly related to nuclear technology but involve sensitive topics, he says.

Mehmet Aksit, a software engineering professor at Twente, worries that the revised policy could toughen an already restrictive visa policy toward Iranians. Although measures to stop nuclear proliferation are appropriate, Aksit says the Netherlands "should encourage intellectual exchanges with Iran."

–YUDHIJIT BHATTACHARJEE

CREDIT: COURTESY OF MANJU RAY

CLINICAL TRIALS

Deaths Prompt a Review of Experimental Probiotic Therapy

The high death rate in a Dutch clinical trial is raising concerns about the use of friendly bacteria, or probiotics, in some patients. Researchers announced last week that in a trial to prevent infections in patients with acute pancreatitis, significantly more patients in the treatment group died than did those in the placebo group. Dutch authorities are now investigating whether the trial design was appropriate and whether probiotics pose any general risks.

In a press conference on 23 January, researchers from Utrecht University in the Netherlands said that 24 patients in the study's treatment arm died after receiving a mix of benign bacteria by feeding tube, compared with nine patients receiving a placebo. That

Difficult news. Researchers Hein Gooszen (*middle*) and Marc Besselink (*right*) discussed the surprising mortality rate in their clinical trial at a press conference.

translates to a mortality rate of 16% for the treatment group, significantly higher than the average expected mortality rate of 10%.

Although it is not clear exactly why the patients died, Hein Gooszen of Utrecht University, one of the study leaders, said at the press conference that it's likely several would have survived if they had not received the bacteria. The results came as a complete surprise, Gooszen said. "There was silence" in the room when the results were unblinded and the difference became clear, he recalled. Although a panel looked at preliminary results of the two groups halfway through the trial, they did not know which had received treatment. That panel found nothing that suggested the trial should be stopped. The researchers say their paper has been accepted for publication but gave no details.

Probiotic bacteria are thought to have a

positive effect on the health of the gut, in part by stimulating the immune system and in part by outcompeting pathogenic bacteria. Strains used as probiotics are typically those that inhabit a healthy gut, such as lactobacilli or bifidobacteria. They have been used to treat a variety of conditions, including allergies and some inflammatory diseases. They are also marketed as a food additive. Two preliminary studies had suggested that probiotics could be beneficial for patients with pancreatitis.

In 2004, the Dutch Acute Pancreatitis Study Group proposed a double-blind trial that would test the use of probiotics in 200 patients. Acute pancreatitis is a sudden inflammation of the pancreas, often seen in patients with alcoholism. A frequent complication is acute infection of the pancreas, which sometimes cannot be treated successfully with antibiotics.

Although most previous trials used just one or two bacterial strains, the Dutch researchers used a mix of six strains; in vitro and animal tests suggested that they had inhibited the growth of the most common pathogens that cause pancreatitis complications. One of the study leaders, Marc Besselink, also of Utrecht University, told *Science* that most patients died of multiple organ failure. He could not comment on whether the multiple strains of bacteria might have set off an unexpected immune reaction.

One explanation for the negative results may be that more severely ill patients ended up in the treatment arm of the study, says Nada Rayes, a surgeon at Humboldt University in Berlin. She says that she and her colleagues have been using probiotics to treat liver-transplant recipients for several years because their clinical studies suggested that it significantly reduced infections and improved outcomes. "I don't want to downplay the results," she says, but in several years of treating immune-compromised, seriously ill patients with probiotics, "we have not had one single extra death."

–GRETCHEN VOGEL

Science Stimulus Unlikely

U.S. science lobbyists hope to persuade Congress to restore some of the money for basic research cut from the 2008 budget last month. But although a few senators, led by Jeff Bingaman (D–NM), are weighing whether to add money to the upcoming Iraq War supplemental appropriations bill, House Speaker Nancy Pelosi (D–CA) says next year is a better bet. "We need to get that money back in '09," she told *Science* last week. "We're all disappointed [with the '08 budget]. But I don't see it getting into the supplemental." Next week's presidential 2009 budget is expected once again to request big increases for the physical sciences. –JEFFREY MERVIS

Ready, Set, Learn

Forget computer bridge and chess. Researchers at the University of St. Andrews in Fife, U.K., are holding a winner-take-all tournament to find the best approaches to solving social-learning problems. The European Commission is putting up €10,000 in prize money for the strategy that proves most effective in computer simulations of situations in which people must compete for resources, such as food. Is it best to copy what everyone else is doing? Rely solely on past experience? Be selective about who to imitate? Entries can be narratives or computer programs and are due 30 June. The organizers will find the 10 best, which will battle each other; the winner will be announced next year. "Hopefully, the field will get a real shot in the arm," says organizer Kevin Laland, a St. Andrews evolutionary biologist. –ELIZABETH PENNISI

Points for Origin-ality

A new virtual center to gather researchers from disciplines related to anthropogeny, the study of human origins, will be announced shortly by the University of California, San Diego (UCSD), and the Salk Institute for Biological Studies. The Center for Academic Research and Training in Anthropogeny (CARTA) is an expansion of a project started more than a decade ago that has quietly gathered leading researchers from diverse fields three times a year to discuss human origins. UCSD biochemist Ajit Varki says $3 million from the Mount Kisco, New York–based Mathers Foundation will support meetings, an archival Web site, a museum, and an eventual journal called *Anthropogeny*. Most centers for human evolution are each "driven by one person focused in their area of research," says Varki, CARTA's organizer. "It's time to start pulling it together." –JON COHEN

CHEMISTRY

DNA Assembles Materials From the Ground Up

For sheer utility, it's hard to beat DNA. It's the molecule of life, of course. And manipulating it is the crux of the entire field of biotechnology, not to mention a novel form of computing. Now, researchers in the United States and Germany report in a trio of papers this week that DNA could also hold the future to building materials from the bottom up.

In two papers in *Nature*, two separate groups in the United States report that they have linked DNA to gold nanoparticles and used it to assemble the particles into extended crystals. Such crystals may prove useful in creating distortion-free lenses and other sought-after byproducts of optical materials known as metamaterials. Down the road, they could also lead to new ways to construct everything from electronic materials to protein crystals coveted by drug developers. The third paper, which appears on page 594 of this issue of *Science*, uses DNA as tweezers to pick up compounds and place them where they're wanted. Although probably too slow for assembling large amounts of materials, the technique could help researchers put chains of molecules together

to answer questions such as how different enzymes work together in a series, each handing off its product to become the starting material of the next enzyme.

Taken together, the results show that nanotechnologists are beginning to look at

Shape shifter. Nanoparticles coated with different DNA sequences can form different crystalline lattices, depending on whether the DNAs bind to themselves (*top*) or to sequences on other particles (*bottom*).

DNA in a whole new light. "This is a watershed in using DNA to assemble objects other than DNA," says John Crocker, a physicist at the University of Pennsylvania.

DNA's power lies in the way chains of its four nucleotide bases pair up, with A's binding selectively to T's and C's binding to G's. As a result, researchers can synthesize single

strands of DNA with a particular sequence of bases, knowing that if the strand encounters its complement the two will bind. More than 10 years ago, teams led by Chad Mirkin at Northwestern University in Evanston, Illinois, and Paul Alivisatos at the University of California, Berkeley, took advantage of this affinity to make DNA sensors. They linked DNA to gold nanoparticles and then added complementary DNA strands to a solution containing the DNA-coated particles. As the DNA partners came together, they drew the particles into an extended network that changed the way light propagated through ▶

DEVELOPMENTAL BIOLOGY

Aging of the Ovary Linked to PTEN Pathway

When a woman is born, her ovaries already contain a full supply of the immature eggs she will need in her reproductive lifetime. Normally, these eggs begin ripening at about age 13 and are gradually released, usually at the rate of one per month, until she is about 50 years old. But in a small minority of women, perhaps 1 in 100, the ovaries stop releasing eggs much earlier in life, thus causing infertility and premature aging. Exactly why that happens isn't understood, but new results may help provide an explanation.

On page 611, a team led by Kui Liu of Umeå University in Sweden, reports that a gene called *Pten*, which is best known as a suppressor of tumor growth, is needed to keep egg development in check. In its absence, the researchers found, the egg-containing follicles of mice were activated rapidly at an early age, thus causing depletion of the animals' eggs much sooner than is normal—a situation similar to that of

premature ovarian failure (POF) in humans.

"It is a very nice piece of work that shows the importance of the PTEN pathway" in controlling follicle maturation in mice, says reproductive geneticist Aleksandar Rajkovic of Baylor College of Medicine in Houston, Texas. If *PTEN* also controls human egg maturation, the finding may aid the design of improved infertility treatments.

Liu and his colleagues came to their conclusion by genetically engineering a mouse strain in which *Pten* expression was inactivated specifically in the animals' oocytes. The results showed that a functional *Pten* in oocytes is needed to keep egg follicles from maturing. Without it, Liu says, "all the primordial follicles were activated prematurely."

Once a follicle is activated, there's no going back. Its egg either matures and is released for fertilization or dies. As a result, the animals had one litter but were infertile by about 3 months of age, which is early

adulthood for mice. By that same age, their ovaries had lost essentially all their follicles.

The result is consistent with previous findings from Diego Castrillon, Ronald DePinho, and their colleagues at Harvard Medical School in Boston. About 5 years ago, they showed that knocking out the gene for the Foxo3a protein produced essentially identical effects (*Science*, 11 July 2003, p. 215). As it happens, *Pten* negatively regulates two other enzymes, PI3 kinase and Akt, which suppress Foxo3a. So with *Pten* gone, Foxo3a can't function.

Castrillon's team, now at the University of Texas Southwestern Medical Center in Dallas, reported in the January issue of *Human Reproduction* that Foxo3a mutations are not a common cause of POF in humans. Still to be determined is whether mutations in PTEN or in other proteins that cooperate with PTEN in producing its cellular effects are involved with human ovarian problems.

Even with that uncertainty, researchers think that the PTEN pathway is a good target for potential fertility treatments. Foxo3a and a number of other proteins shown to affect egg development in

the solution and produced a color change (*Science*, 22 August 1997, p. 1036). Today, the technique is at the heart of DNA sensors now on the market.

But although the early effort showed it was possible to make clumps of DNA-linked particles, it couldn't organize those particles into regular crystalline arrays. Crocker and Mirkin say they think numerous groups tried and failed to find the magic recipe.

In *Nature*, Mirkin and colleagues, as well as a second group led by Oleg Gang at Brookhaven National Laboratory in Upton, New York, report hitting on a similar solution: longer DNAs. The trouble with previous efforts, Gang says, is that the DNA linkers they used between nanoparticles were likely too short and thus relatively rigid. By substituting longer, floppier DNA molecules, Mirkin's and Gang's groups gave the particles room to move around and settle into a conformation that maximized the number of links they could make with their neighbors. The result was that both groups were able to program their materials to form what is known as a body-centered cubic arrangement, a crystalline lattice in which two particle types alternate their positions at the corners and the center of the cube (see figure, p. 558). And both Gang and Mirkin expect this is only the beginning. "There is a

hope that many more structures will be possible," Gang says.

That hope is echoed by Hermann Gaub, whose team at the University of Munich, Germany, also pushed the limits on using DNA to organize other materials. Gaub's team used DNA as a tool to assemble materials molecule by molecule. A version of this experiment was first reported in 1990, when Donald Eigler and colleagues at IBM used the tip of a scanning tunneling microscope to assemble xenon atoms on a surface to spell "IBM." But that required a vacuum chamber and a temperature near absolute zero.

Gaub and colleagues essentially repeated the feat at room temperature and in water, using as building blocks light-emitting molecules known as fluorophores. First, they linked the fluorophores to DNA strands; then, using complementary strands attached to the tip of an atomic force microscope, they moved them one by one onto a surface where other DNA molecules held them in place. Gaub says he hopes the technique will help engineers further miniaturize biomedical devices by allowing biomolecules to be spotted down with exquisite control. Such specificity is DNA's forte and could help the biomolecule gain a new job as a construction tool for the nanoworld.

–ROBERT F. SERVICE

Tissues Case Over

Participants who have donated tissue for research shouldn't get their hopes up that they can ever take it back. That issue was largely settled last week when the U.S. Supreme Court declined to intervene in a fight between a medical researcher and his former university over who owns tissue donated by patients. William Catalona, a prostate cancer researcher at Washington University in St. Louis, Missouri, sued the school after it blocked him from taking his patients' tissue samples to a new job at Northwestern University in Evanston, Illinois.

The university argued successfully in federal district and appellate courts that institutions own the biological samples their researchers collect because patients have donated them as gifts (*Science*, 29 June 2007, p. 1829). Law professor and bioethicist R. Alta Charo of the University of Wisconsin, Madison, says that although the district court's ruling in favor of the school may not apply in every case—it depends on a particular state's gift law and the details of the patient consent form—"the opinion is likely to be highly influential nationwide." Catalona says that the legal rulings have made "a travesty "of federal regulations protecting research subjects.

–JOCELYN KAISER

Running out. Mouse ovarian tissue in which *Pten* was inactivated in oocytes (*left*) shows many activated follicles just 8 days after birth. But by 12 weeks (*right*), the mouse ovaries are depleted of follicles.

animals are transcription factors that regulate gene expression. But PTEN is an enzyme, and developing compounds that inhibit it, thus promoting egg development and maturation, should be easier. An enzyme "has the advantage of being easily manipulated," says Aaron Hsueh, an ovarian physiologist at Stanford University School of Medicine in Palo Alto, California.

He notes that although most women with POF are diagnosed after all their eggs are depleted, some still have follicles in

their ovaries. It's impossible to tell when they will mature and release an egg, but treatment with a PTEN inhibitor could allow physicians to control that and perhaps help a woman become pregnant. Another possibility is to use a PTEN pathway inhibitor to aid follicle maturation in ovarian tissue in lab dishes, as may be needed for women who have had their ovaries removed prior to chemotherapy that would render them sterile.

–JEAN MARX

EPA Wants Data: If You Please

How risky are nanomaterials? Last week, the U.S. Environmental Protection Agency (EPA) launched a voluntary nanomaterials toxicity reporting program to help find out. The new program encourages—although it does not mandate—companies that make, use, or import nanomaterials to submit characterization and risk data to help the agency figure out which nanomaterials are worrisome.

The agency has also affirmed its previous decision that it will not consider nanoparticles new chemicals if they are made of the same chemicals as materials currently registered under toxicity databases. The rule would apply, for example, to nanoparticles of zinc oxide because larger clumps of the stuff are a component of skin creams.

That's a mistake, says Andrew Maynard, chief science adviser to the Woodrow Wilson International Center for Scholars Project on Emerging Nanotechnologies. "EPA's approach ignores the existing scientific research that suggests different nanostructures with the same molecular identity present different hazards," Maynard says.

–ROBERT F. SERVICE

The answer? Much work remains to be done to prove iPS cells.

A Seismic Shift for Stem Cell Research

The development of pluripotent cells from individual skin cells has opened up a new world of research, but scientists say they still need to work with embryonic stem cells

Advocates and opponents of embryonic stem (ES) cell research don't agree on very much, but both have long said that avoiding the use of human embryos—through direct reprogramming of cells—would solve a lot of problems, both ethical and scientific. Now that goal has been achieved, but the picture is not quite as simple as some had hoped.

At the end of last year, two teams—one in Japan, the other in the United States—announced that they had generated human ES-like cells simply by introducing a handful of genes into skin cells. Soon after, a third group at Harvard University joined the fold, also reporting the creation of these cells, called induced pluripotent stem (iPS) cells, from a variety of tissues.

But after the initial excitement, both sides are finding that although iPS cells have answered some questions, they have also raised new ones. Opponents of ES cell research were quick to argue that embryos are no longer needed to realize the promise of stem cells and that tight restrictions on ES cell research should be maintained, even strengthened. But researchers were equally quick to respond that it is too early to close the book on human ES cells, especially because it is not clear that iPS cells will ever be safe for use in patients.

New promise

Shinya Yamanaka of Kyoto University in Japan sent a wave of excitement through the stem cell field when he announced in June 2006 that he had reprogrammed mouse skin cells into something that closely resembles ES cells by inserting just four functioning genes into the cells (see sidebar, p. 562). The cloning experiments that produced Dolly the sheep—which involved removing the nucleus of an egg cell and replacing it with one from an adult cell—had demonstrated that the DNA in adult cells is still capable of directing development, when given the right cues. But no one knew whether those cues came from hundreds of genes or just a few. Yamanaka's work showed that a handful could do the trick, suddenly turning the idea of embryo-free pluripotent cells into reality.

When Yamanaka's group, and a group led by James Thomson of the University of Wisconsin, Madison, announced in November 2007 that they had performed the same feat with human cells, the excitement turned into a frenzy. Research teams around the world are now adding iPS cells to their repertoire. "It is an incredibly rapidly moving field," says Harvard stem cell researcher George Daley, author of the most recent iPS paper. "There's a huge energy to the science, a sense of infinite possibility." Cultivating iPS cells is "so easy to do now, papers are coming out every week," adds Konrad Hochedlinger of the Harvard Stem Cell Institute in Cambridge, who predicts "a huge gusher of new data."

Because iPS techniques offer a way to generate cell lines from a patient's own tissues, they supply "an incredibly powerful new tool" for research, says stem cell researcher Renee Reijo Pera of Stanford University School of Medicine in Palo Alto, California. Previously, scientists had thought the only way to create a cell line from a patient with, say, Parkinson's disease was through the technique that produced Dolly, known as somatic cell nuclear transfer (SCNT) (see diagram, p. 561). That requires hard-to-obtain human oocytes and involves the destruction of an embryo. But now, scientists can reprogram cells with tools available in any molecular biology laboratory. "It's like transfection. It's really very straightforward," says Hans Schöler of the Max Planck Institute of Molecular Biomedicine in Münster, Germany.

Kevin Eggan's group at Harvard is already collecting skin cells from patients with amyotrophic lateral sclerosis to generate iPS cell lines. Similarly, Lawrence Goldstein, director of the stem cell program at the University of California, San Diego, says his group is about to collect skin biopsies from Alzheimer's patients. In the meantime, he says, he has put plans for SCNT research on hold. The iPS lines will be useful not only for studying disease mechanisms, he says, but also for testing the efficacy of new drugs against those diseases.

CREDIT: U. OF WISCONSIN

IPS cells may also help speed the search for better ways to coax pluripotent cells to differentiate into desired cell types. Currently, scientists can reliably steer human ES cells to make heart cells and several kinds of neurons, but reproducible chemical recipes for most cell types are still elusive.

Embryo-free, eventually?

The promise of iPS cells has prompted some opponents to assert that human ES cell research is no longer necessary. "The embryonic stem cell debate is over," wrote columnist and physician Charles Krauthammer, a former member of the President's Council on Bioethics, in a widely cited 30 November 2007 column. "Scientific reasons alone will now incline even the most willful researchers to leave the human embryo alone."

But that's far from the case, scientists say. "We'd be glad to eventually give [embryo research] up," says Douglas Melton, a Harvard stem cell biologist. But "it would be premature" to do so now.

At every step in iPS research, comparisons with ES cells will be required, Daley points out. "Right now, we're not certain iPS cells are the absolute equivalent" of ES cells, capable of forming all the desired tissue types. Adds Fred Gage of the Salk Institute for Biological Studies in San Diego, California: "Many of the existing ES cell lines are different from each other, so it depends what you compare them to. It's likely that iPS lines are different from each other" as well. To validate and improve iPS cells, says Eggan, scientists will "need to make huge [numbers] of these cells from many different people and compare them in a battery of tests with human ES cells." This will take a while because scientists can't do definitive experiments with human cells that they can do in mice, such as creating chimeric animals.

But even then, some scientists are dubious that iPS cells will ever substitute for ES cells in therapies for heart, neurological, and other diseases. The reason: So far, scientists have used retroviruses to ferry the reprogramming genes into the cells, and those viruses may interfere with important genes and lead to cancer. Right now, says Eggan, there's "no clear road to getting rid" of those retroviral vectors.

Harvard's Hochedlinger says one alternative might be a "transient delivery system" such as adenoviruses that don't permanently insert into a chromosome. But the best option would be to forgo introducing genes and

To validate iPS cells, scientists must "make huge [numbers] … from many different people and compare them in a battery of tests with human ES cells."

—KEVIN EGGAN, HARVARD UNIVERSITY

instead use small molecules to slip through a cell's membranes and into the nucleus to turn on the genes that make the cell revert to a pluripotent state. Biologist Ding Sheng of The Scripps Research Institute in San Diego says he is partway there. In as-yet-unpublished work, Ding says his group has found small molecules that will substitute for some of the genes needed to turn mouse cells from a variety of sources—including neurons and skin cells—into iPS cells. "We start by throwing in all four genes and seeing which molecule can improve the process, then we start subtracting genes," he says. Advanced Cell Technology (ACT) in Worcester, Massachusetts—a U.S. company with a large commitment to ES cells—is also at work on finding ways to turn on the relevant

genes in adult cells without inserting either viruses or genes, says Robert Lanza, chief scientific officer at ACT.

But at least one company that hopes to commercialize stem cell treatments says it is sticking with human ES cells. Geron Corp. in Menlo Park, California, which funded the initial research on deriving human ES cells and which owns key licenses and patents involving their therapeutic use, has "no plans to deviate from naturally derived ES cells," says company president Thomas Okarma. He says that's based not on intellectual property claims—he asserts that Geron's licenses apply to all pluripotent cells, including iPS cells—but on science. Even if iPS cells can be grown without the aid of a potentially cancer-causing virus, he says that they "can't possibly be used for therapies." Starting with a skin cell that might have been altered by aging or toxins instead of a "pure crystal-clear" human embryo would add unpredictable risks, he explains.

Even if iPS cells eventually prove safe for use in humans, scientists say the notion of

Cloning (SCNT)

Induced pluripotent stem cells

Donor skin cells

Skin cell fused with enucleated egg

Retroviral infection with four genes

Skin cell nucleus reprogrammed

The same? There are now two ways to make genetically tailored stem cells.

Blastocyst with inner cell mass

ES cells

Pluripotent cells

Possible iPS colonies
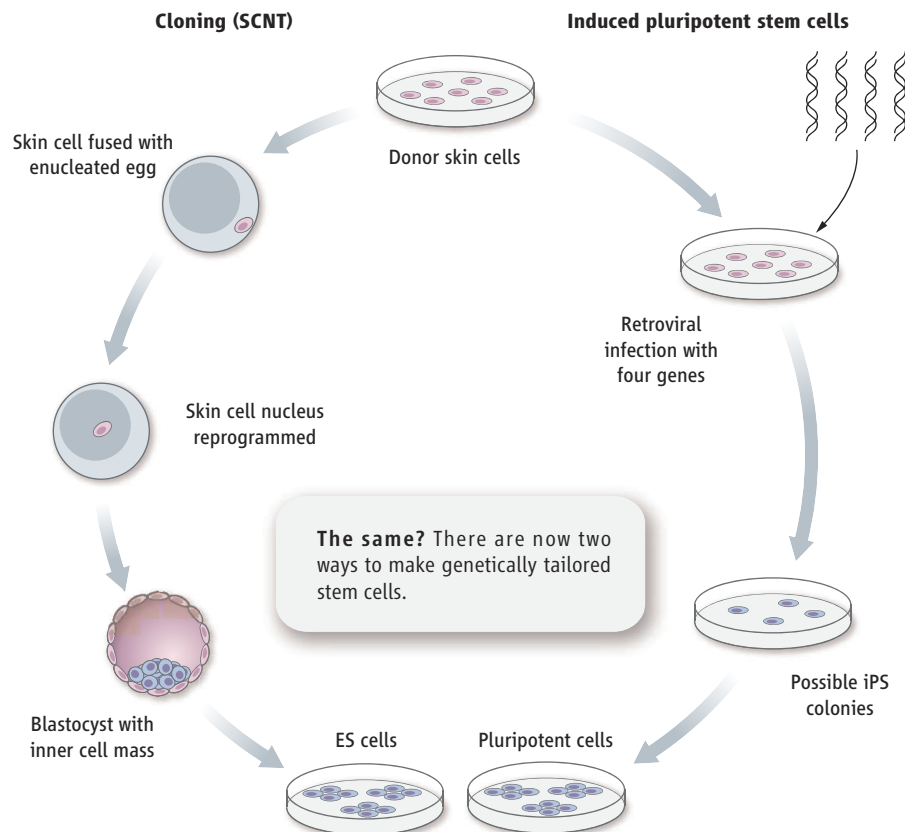
SHINYA YAMANAKA: MODEST RESEARCHER, RESULTS TO BRAG ABOUT

Few researchers have rocketed from relative obscurity to superstar status as quickly as Shinya Yamanaka. At the beginning of 2006, his work on stem cells was little noted beyond Japan's Kyoto University, where he has been a professor since 2004. But he gained international scientific recognition in mid-2006 when he reprogrammed adult mouse cells to behave like embryonic stem cells without the use of embryos. Then, last November, when he showed he could replicate the trick using adult human cells, he garnered mainstream press coverage around the world, earning praise from bioethicists who object to research involving human embryos and generating whispers of a possible Nobel Prize in the scientific community.

Through it all, the man at the center of this attention maintains it was almost serendipity. "We were extremely lucky," Yamanaka says. "I know many other scientists who have been working harder and who are smarter than we are."

Yamanaka's colleagues disagree. "It's a very original break-through," says Norio Nakatsuji, a stem cell researcher and colleague at Kyoto University. "What Yamanaka did is really spectacular," says Douglas Melton, a stem cell researcher at Harvard University. "He deserves an enormous amount of credit for the conception of an approach that really paid off," says Melton. Several scientists say the work is "Nobel-worthy." Characteristically, Yamanaka demurs when the "N" word comes up.

Yamanaka's modesty and politeness have contributed to his celebrity in Japan, where his story has been told in countless newspaper and magazine articles. A medical doctor who trained to be an orthopedic surgeon, Yamanaka, 45, switched to research to explore new medical treatments. He was inspired to search for alternatives to using human embryos to create stem cells when a peek at an embryo under a microscope at a friend's fertility clinic reminded him of his two daughters as infants. He identified genes active in early mouse embryos and then painstakingly tried them one by one and then in combinations to see if they would turn adult mouse cells back into stem cells. He was surprised that just four genes could reprogram adult mouse cells, as he announced at the International Society for Stem Cell Research conference in Toronto, Canada, in June 2006. And he was further surprised when the same four genes reprogrammed adult skin cells as well. "We thought making human induced pluripotent stem [iPS] cells would be more difficult than it was," he says.

Avoiding the use of embryos to produce stem cells promises to open up the field in countries where work had been restricted or banned on ethical grounds, although Yamanaka himself emphasizes that work on human embryonic stem (hES) cells must continue. The liberating effect may be particularly important in Japan. Nakatsuji, who heads the only group in Japan to derive hES cells, says studies using the cell lines have been slowed here by stringent bioethical oversight requirements and other hurdles. Work on iPS cells is not likely to be as strictly regulated, he says.

Already, Japan's Ministry of Education included a $21 million package of programs for iPS research in the budget for the fiscal year starting 1 April. A portion of the money will go to support the new iPS Cell Research Center at Kyoto University, which Yamanaka will direct. Another portion will support grants for iPS cell work.

Yamanaka points out that although the budget for iPS cell work is large by Japanese standards, it is still only about one-tenth of the amount available annually in the United States. In addition, in the United States "there are all sorts of researchers with differing expertise within one institution; that's a huge advantage," he says.

A different path. An original approach to generating stem cells has made Shinya Yamanaka a celebrity.

To bring together Japan's dispersed scientists, the ministry is planning to create an iPS cell research consortium that will try to iron out material-transfer and intellectual-property issues and investigate other ways of supporting collaborations. "It will be something like a virtual institution," says Yutaka Hishiyama, director of the ministry's Life Science Division.

Yamanaka's focus will remain on creating and studying iPS cells, leaving their differentiation into the various cell lineages to others. His group found that less than 1% of all cells transfected with the four genes became iPS cells. He is mulling over how to study what's going on in this "black box" so as to improve the efficiency and possibly identify better ways of creating iPS cells. The group already reported that it has reduced to three the number of genes required to generate an iPS cell. For now, the team is continuing to use potentially cancer-causing retroviruses to carry the genes into the cells, which could make the cells unsuitable for clinical treatments. To circumvent this problem, the group is investigating purely chemical approaches as well. Yamanaka also wants to determine the safety of iPS cells made both with and without retroviruses, something likely to require long-term animal studies. He suspects the first clinical uses may be a decade away. "But for other applications, like toxicology and drug discovery, [iPS technology] is ready to go," he says.

Last August, Yamanaka set up a small lab at the Gladstone Institute of Cardiovascular Disease in San Francisco, California, where he had done a postdoc in the 1990s. He originally thought of moving there full-time to sidestep Japan's restrictions on using hES cells. But quicker-than-expected success in creating human iPS cells has made that unnecessary, he says. Although he intends to keep the Gladstone connection, he says he now feels an obligation to remain in Japan. The tax money being dedicated to iPS research makes for "a huge responsibility," he says, and public interest creates "a great chance to make the Japanese science environment more comfortable." Deep down, however, he says, "I hope [the fuss] will quiet down after 3 or 4 months so I can get back to research."

–**DENNIS NORMILE**

With reporting by Gretchen Vogel.

generating individually tailored cell populations for every patient will still be a pipe dream. "Patient-specific therapy is totally impractical, even with iPS," says Lanza. "You would need millions of lines." Furthermore, there's no time to generate the cells if they are needed rapidly, as after a heart attack or spinal injury. "The idea that you would use iPS cells for individual treatments is lunacy," agrees Stephen Minger of King's College London. "It takes 6 months of really hard slog to make a cell line."

Instead, Lanza and others say, the most likely approach will be to create banks of 100 or so cell lines with different immune properties that would provide acceptable matches with most of the population. They could be generated from donated embryos, iPS cells, or cell lines gained through SCNT. Because no human cell lines have yet been generated via SCNT, it's far too early to know which would work better, he says (see sidebar, right).

Political question marks

For that reason, iPS cells have sharpened rather than solved the ongoing political battles over embryo research. "For once," says Eggan, "the cell du jour for the opponents is actually based on good science"—and that has only reinforced the opposition. In the United Kingdom, the Human Fertilisation and Embryology Authority postponed a decision on whether scientists should be allowed to insert DNA from human cells into cow eggs, in part to consider the implications of iPS cells. (It gave Minger and another group the green light in January.) Opponents also cited iPS cells in their efforts to add new restrictions to a long-planned law updating regulations on embryo research.

In the United States as well, the cells are influencing several ongoing debates. In Missouri, voters in late 2006 amended their constitution to allow any type of embryo research not banned by federal laws—including SCNT. But iPS cells are bolstering opposition to the

IPS cells "can't possibly be used for therapies."

—THOMAS OKARMA, GERON CORP.

NUCLEAR TRANSFER: STILL ON THE TABLE

With so many unknowns, scientists want all research avenues kept open. So somatic cell nuclear transfer (SCNT) research continues apace. Now that money is flowing from the California Institute for Regenerative Medicine, several new groups are gearing up and others are continuing. For instance, Renee Reijo Pera of Stanford says that SCNT "remains the only tool" for her purpose, which is to learn how to bypass defects in egg reprogramming in order to help infertile women.

In Massachusetts, two groups at Harvard remain intent on generating cell lines from cloned embryos. "SCNT is a tool for asking questions about early development that won't be addressed by iPS [induced pluripotent stem] cells," says George Daley, whose group is experimenting with leftover eggs from fertility clinics. What's more, he says, the process can supply insights into reprogramming that could be "very, very important for improving the fidelity of iPS reprogramming."

Obtaining eggs for this purpose remains a problem. Fresh eggs, not fertility-clinic discards, seem to offer the best chance for successful generation of cell lines from SCNT. In the United Kingdom, researchers at the University of Newcastle have managed to recruit 15 oocyte donors by offering to cover half of the costs of their fertility treatment in return for the donation of half of the oocytes retrieved during the treatments. And in Spain last week, Miodrag Stojkovic of the Centro de Investigación Príncipe Felipe in Valencia received a license to pursue his experiments, also using leftover oocytes from fertility treatments.

But at Harvard, a group led by Kevin Eggan and Douglas Melton has had "terrible" results in attempting to get egg donations from young women. "After almost 2 years, … more than $50,000 in advertising, and many hundreds of calls from interested women," there hasn't been a single donation. The reason, says Eggan: "They know they can get paid for doing exactly the same thing for assisted reproduction. So why do it for free?"

A close relationship with a fertility clinic seems to have paid off for a San Diego, California, company. Stemagen announced last month that it has derived at least two human blastocysts (see photo, above) from inserting the nuclei of skin cells into 25 fresh eggs donated by women aged 20 to 24 who had served as successful egg donors for infertile couples and had spare oocytes left over from the procedure.

—C.H. AND G.V.
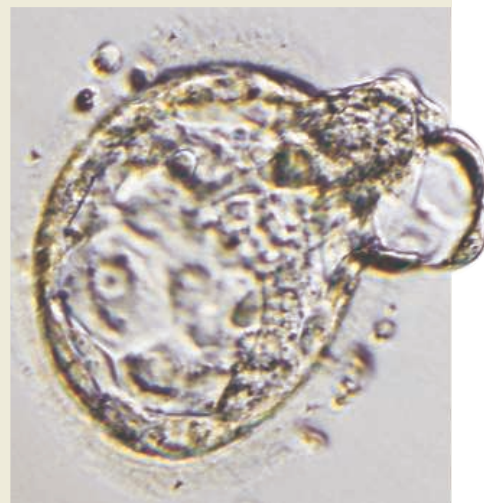
measure. Opponents have regrouped and are pushing to put the question back on the ballot in November. Now, says William Neaves, president of the Stowers Institute for Medical Research in Kansas City, "efforts of opponents are so overreaching and broad, they could affect even [research using] existing [ES cell] lines."

On the national level, an aide to Representative Michael Castle (R–DE) says that Castle intends to introduce legislation again in the next Congress to overturn the Bush Administration's ban on federal funding for work on ES cell lines derived since August 2001. Similar bills sponsored by Castle and Representative Diana DeGette (D–CO) were twice vetoed by President George W. Bush. The cloning wars may also be starting up again on Capitol Hill: Representative

Dave Weldon (R–FL) has announced plans to reintroduce a cloning ban that would include SCNT, and Senator Dianne Feinstein (D–CA) says she will reintroduce a bill that bans reproductive cloning but allows SCNT. "The iPS developments will require additional educational efforts" on Capitol Hill, predicts lobbyist Tony Mazzaschi of the Association of American Medical Colleges in Washington, D.C.

Where the political debate will go next is anyone's guess, says stem cell researcher Sean Morrison of the University of Michigan, Ann Arbor. "Nobody could have predicted the twists and turns this field has taken over the past 3 or 4 years. … If this experience shows us anything, it is that we can't predict where the field is going to be even 1 year down the road."

—CONSTANCE HOLDEN AND GRETCHEN VOGEL

NASA

Scientists Hope to Adjust the President's Vision for Space

An inadequate budget and daunting technical challenges will force the next U.S. president to rethink current plans for a postshuttle NASA. Space scientists are offering input on what those changes might look like

When U.S. President George W. Bush laid out his plan for a revamped civilian space program in January 2004, he said it would provide "a great and unifying mission for NASA." That expansive vision included a launcher to replace the shuttle, a lunar base, and a slew of robotic missions to the moon and Mars that would put smiles on the faces of even the most skeptical planetary scientists.

But 4 years later, that vision has instead triggered a civil war among competing interests within the space community. Some space researchers want to delay the launcher and a lunar base to protect the stalled science budget, whereas industry lobbyists are pressing hard to speed up those schedules. Both groups worry that the next Administration will want to spend more to monitor Earth at the expense of robotic or human exploration of the moon. "Everyone agrees the [current] program is too ambitious, that there is not enough money, and that we are killing earth sciences while eating our technological seed corn," says space economist Molly Macauley of Resources for the Future in Washington, D.C., who attended a recent forum in Irvine, California, on the future of the civilian space effort.

The late November forum, sponsored by the U.S. National Academies, marked the start of a new community effort to rescue a troubled agency. This month, two more meetings with a similar agenda will take place, one at Stanford University in Palo Alto, California, and the other at a Washington, D.C., think tank. Among the options participants will consider are simplifying the shuttle replacement, delaying plans for a lunar base and focusing on an asteroid or the moons of Mars as a better destination, and teaming with other countries to defray the cost of expensive missions such as a Mars sample return. The organizers of each meeting hope to infuse the current presidential campaigns with their findings and influence the policies of the winner in November (see sidebar, p. 565). "There has to be a rethinking," says Lennard Fisk, chair of the Space Studies Board under the academies' National Research Council. "We're not on a sustainable course."

The crisis stems from a slew of unexpected financial setbacks in recent years. Finishing the space station by 2010 and retiring the shuttle that same year will cost many billions of dollars more than anticipated. The new Ares launcher, scheduled for its first flight in 2014, already is encountering technical difficulties, and its cost is certain to climb. And ballooning price tags for many research projects in recent years—such as a $1.5 billion increase in the James Webb Space Telescope (JWST) alone—has put a further strain on NASA's $17.3 billion budget. "The NASA budget is about $2 billion to $3 billion short," says Fisk. "Cannibalizing other parts of the NASA program to make even limited progress in building Ares and Orion [its capsule] is not a strategy for success but rather one of desperation and will not succeed."

Neither the White House nor Congress seems willing to provide the extra dollars needed. And nobody expects the next Administration to push for dramatic increases, either, except perhaps in the earth sciences. As a result, "the NASA program is tremendously unstable," warns Charles Kennel, former chair of the NASA Advisory Council, who attended the Irvine meeting. "And it is going to stay so until the shuttle is definitively retired."

Shortsighted vision

It wasn't supposed to be this way. The bulk of the money for the new vision that President Bush laid out at NASA headquarters (*Science*, 23 January 2004, p. 444) was to come from retiring the expensive shuttle system. And the overall annual budget for space science was slated to grow from $3.9 billion to $5.6 billion to ramp up projects such as a Mars sample return by 2013. That plan quickly unraveled with the unexpected overruns for science and the shuttle, which required billions in repairs to the fleet after the disintegration of Columbia on 1 February 2003.

In response, NASA Administrator Michael Griffin, who succeeded Sean O'Keefe in April 2005, diverted $3 billion projected for space science through 2011 to cover the added shuttle costs. He halted work on a host of astrophysics and earth sciences projects and abandoned the series of robotic lunar landers. In addition, Griffin slashed funding for aeronautics research and life and microgravity sciences. "It was a $4 billion problem, and it was painful to fix," Griffin told the American Astronomical Society (AAS) at its January meeting in Austin, Texas.

Scientists are still feeling the pain. At the Irvine meeting, there was "discouraging pessimism about the way the space program has fared in the past 4 years," says Fisk. And few

WHAT'S LEFT OF A SOARING VISION ...

Aborted launch? NASA is moving ahead with the Ares rocket, which will replace the shuttle. But much of the robotic portion of Bush's 2004 vision has been jettisoned.

2008: Lunar robotic orbiter
2009: Annual lunar robotic landers begin
2011: Mars demo lander mission
2013: Mars robotic sample return
2014: Test flight of shuttle replacement
2015: Nuclear power/propulsion demo
2018: Astronauts return to the moon

CREDIT: NASA

Launch control. Obama has reworked his position on a shuttle replacement.

Getting Up to Speed on Space

It was a single sentence in the last paragraph of a 15-page white paper on education reform. But when presidential candidate Senator Barack Obama (D–IL) suggested in November that he would partially finance those reforms by delaying NASA's new launch vehicle for 5 years, space enthusiasts let out a howl. "It was not full-fledged blowback," says one senior Obama staffer who requested anonymity. "But we did hear from people who wanted an explanation."

Last month, shortly after Obama's victory in the Iowa caucuses, they got one. The candidate's first position paper on space not only promises to stick to the current schedule during at least the early phase of the launcher, but it also backs "a bold array of robotic missions" and pledges a "much-needed infusion of funds" for federally funded scientific research.

What happened? Obama's sudden support for the rocket that will replace the shuttle as the country's new vehicle to explore space demonstrates the surprising success of grassroots and Internet-based space efforts in affecting the course of the 2008 presidential campaign. Blogs such as Space Politics and NASA Watch, and organizations such as the Mars Society, keep a close eye on every utterance by a candidate on space policy. They instruct their audience how to contact the campaigns and even coach readers on how to get a space question inserted into a presidential debate. And they are being heard.

My space. Exploration advocates are using the Web as an organizing tool to influence the 2008 presidential elections.

"It's a small but vocal group, and they've reached out from the beginning," says the Obama staffer. "I'm impressed with the grassroots effort," adds Lori Garver, a Washington, D.C., space consultant and former NASA official who advises Obama's chief rival, Senator Hillary Clinton (D–NY). "They've done more than all the sophisticated lobbyists."

Space has never been on the radar of so many candidates so early in a presidential campaign. Its presence is forcing campaign staffs to be familiar with acronyms, engineering plans, and flight timetables to a degree

unthinkable in previous elections. Clinton, for example, promised in October to speed up construction of the new rocket, restore cuts to the agency's aeronautics effort, and pour new money into a space-based climate change initiative. "This is not a traditional core issue for a campaign," admits a senior staffer about the increased attention to space policy. "It's a pretty steep learning curve."

The increased scrutiny may also be the reason we already know that the sharpest difference between Clinton and Obama on space involves not whether to build the Ares rocket but where it should go once it's ready for launch. Neither candidate specifically backs President George W. Bush's goal of sending astronauts back to the moon by 2020, but a senior Clinton staffer said last week that she "will support future missions to the moon." Obama, by contrast, is not yet sold on a lunar base as a sensible or necessary step. His campaign staffer predicts that the "later phases" of NASA's exploration plans will be delayed.

Some policy analysts say that the rising interest in space is also being driven by old-fashioned worries about jobs. "Candidates feel compelled to talk about NASA not because it's a topic of intrinsic interest," says David Goldston, a former House Republican staffer who is now a visiting lecturer at Harvard University. "At some point it becomes an important local issue they can't ignore. They are stuck with it rather than drawn to it."

The economic factor is heightened by the fact that the space shuttle, the mainstay for U.S. human exploration over nearly 3 decades, is finally being put out to pasture as of 2010. Its retirement threatens to lay off thousands of workers in the politically crucial state of Florida, home to the agency's sprawling Kennedy Space Center. The projected 4-year time gap between flying the shuttle and a new rocket is another hot local issue.

Accordingly, Republican presidential candidates Rudy Giuliani and Mitt Romney made separate pilgrimages in the days before the state's 29 January primary to tour shuttle facilities, shake hands with workers, and assure business leaders that they won't abandon an area dubbed the Space Coast. "Florida will continue to be the center of America's space program ... and our emerging space industry," Giuliani promised a group of local business executives, whereas Romney spoke warmly of Bush's push to return to the moon (see main text). Another Republican presidential contender, veteran Senator John McCain (R–AZ), declined the group's invitation. But McCain already knows the players in the space community, thanks to a stint as chair of the Senate committee that oversees NASA.

Campaign staffers say that they welcome input beyond the usual circle of lobbyists on how to shape science, space, and technology policies. "We're trying to spread a wider net," says the Obama staffer. That gives researchers and engineers a new opportunity to plug into the political process. "They are looking for politically seasoned people they can trust," Garver says about the candidates. "And now it's so easy to have an impact." **–A.L.**

expect immediate relief. Specifically, the cost of completing the space station, retiring the shuttle, and building the new launcher continues to climb. And there is little prospect that any new president or Congress would cancel the human-exploration effort. "Nobody wants to be the first to say there should be no human space flight program" for fear of sparking a public outcry, says David Goldston, formerly a Republican House staffer and now a visiting lecturer at Harvard University. Some lawmakers are already nervous about the lengthy delay—4 years and counting—between shuttle retirement and the introduction of the new rocket. "If anything, the political pressure to decrease the gap will put increasing pressure on the science budget," says George Whitesides, executive director of the National Space Society in Washington, D.C.

Moon or bust?

In his AAS speech, Griffin urged scientists to support the goals of the human-exploration community. But many researchers argue that stretching out the human-exploration timetable would ease the financial stress on NASA without damaging the program. "The date for a return to the moon may be overambitious," says Kennel. "Let's be realistic; if it slips, it slips." Wesley Huntress Jr., a former NASA science chief and current president of the Planetary Society in Pasadena, California, is blunter: "Go to the moon by 2020? It ain't going to happen, so it is kind of silly to keep talking about it."

Huntress's organization is cohosting the meeting at Stanford to consider alternatives to the current program. Stanford astronomer G. Scott Hubbard, also a former senior NASA manager, says the closed gathering of engineers and scientists hopes to come up with clear goals that balance the needs of exploration and science. "Let's find common ground and stop heaving grenades," he says.

Among the options to be discussed are human missions to an asteroid or to the martian moons, Phobos and Deimos. In order to reach an asteroid, spacecraft need less fuel than is needed to reach a planet because an asteroid's weaker gravitational field means that less energy is required to reach the surface. Former astronaut Russell "Rusty" Schweickart notes that asteroids are "a much more rational choice" as a destination because they are richer in resources that could potentially be mined, and because they could pose a deadly threat to Earth.

Scientists say a successful mission will do more than increase their knowledge of how the solar system evolved. "People could be tremendously excited by it," says Whitesides. "The moon was the logical steppingstone a
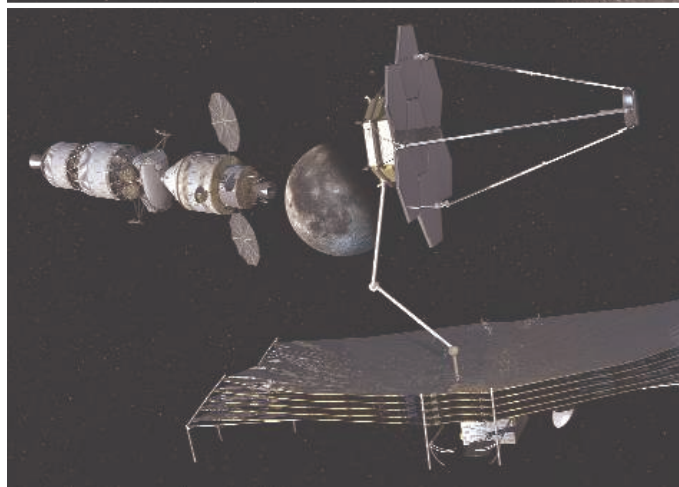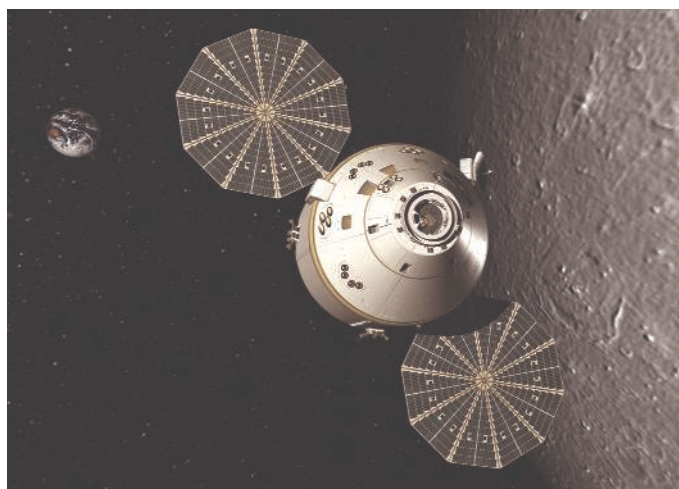
Moonbound? NASA's new launcher is now focused on a lunar return (*above*), but researchers would also like it to be capable of assembling large telescopes in space.

half-century ago, but what's changed since is our understanding of the importance of near-Earth objects." But that redirection may be hard to sell to many U.S. policymakers, who worry that such a program would pale in the public eye if China succeeds in its announced goal of sending humans to the moon.

The Stanford meeting will also review the current plans for Ares, asking if it would be cheaper to modify an existing rocket and how the new launcher could accommodate the needs of scientists as well as astronauts. Astronomer Alan Dressler of the Observatories of the Carnegie Institution of Washington in Pasadena, who took part in the Irvine meeting, would like to see a launcher capable of building and servicing large-scale telescopes in deep space. That would require an airlock, which is not budgeted in the current program.

NASA officials, however, insist that Ares and the Orion capsule that eventually will sit on top of it will play a key role in advancing science as well as human exploration. "Given foldable optics, you could launch a telescope two to three times as big as JWST," says Edward Weiler, director of NASA's Goddard Space Flight Center in Greenbelt, Maryland.

Renewed interest in international cooperation could also rescue long-delayed missions. For example, NASA is now considering a joint multibillion-dollar Mars sample-return mission with the European Space Agency, and the next president might want to invite India or China to join the space station effort. "A new Administration is likely to use exploration as a catalyst for international cooperation," says retired ambassador Roger Harrison of the U.S. Air Force Academy in Colorado Springs, Colorado, which this month is cohosting a gathering on the future of space with the Center for Strategic and International Studies in Washington, D.C.. "We need to start thinking like the particle physicists," adds Kennel, pointing to international facilities such as CERN. "Some programs are beyond the capacity of the United States."

Even if the new launcher is delayed and other countries can shoulder some of the cost, nobody expects the fiscal problems dogging space and planetary sciences to evaporate when the next president takes office. NASA's budget is likely to remain flat for the foreseeable future, says Griffin: "We would do best to plan accordingly." Revitalizing NASA's earth sciences programs is a high priority for several candidates because of the growing public concern about climate change. And pressure to hold down domestic spending because of the cost of the Iraq war and a growing budget deficit makes it unlikely that NASA's overall pool of funding will rise sufficiently to raise all boats. Says Huntress: "We're going to have to live with less."

–ANDREW LAWLER

The blue pearl. A yak searches for vegetation near Lake Hovsgol.

The Big Thaw Reaches Mongolia's Pristine North

As warmer temperatures affect plants, animals, and human society, researchers ask whether ecological changes can be reversed

LAKE HOVSGOL, MONGOLIA—The shortcut through the forest was the only way to get back to base camp before dark. But now Tumorsuk and his passengers may not make it home tonight at all. The burly forest ranger steps on the gas pedal, and the wheels scream as the jeep sinks deeper into the muck. Tumorsuk mutters in Mongolian, kills the engine, and climbs out. The fading daylight casts a blue hue on fresh snow covering every pit, branch, and boulder. Wolves will soon be on the prowl.

"He wants us to get out," ecologist Bazartseren Boldgiv says calmly. He and his Ph.D. student Lkhagva Ariuntsetseg, both at the National University of Mongolia in Ulaanbaatar, peer under the jeep; its belly is flush with the ground. Ecologist Clyde Goulden, visiting from the Academy of Natural Sciences in Philadelphia, Pennsylvania, scans the evergreen taiga forest. The trees tilt at drunken angles. Some have toppled over. "This is climate change," Goulden says. Higher average temperatures in summer are thawing the layer of permanently frozen soil, or permafrost, and disturbing the soil structure around the shallow tree roots.

Global warming is not a uniform process. Mongolia, particularly at the high altitudes around Lake Hovsgol, has been warming more than twice as fast as the global average. Unique ecosystems are feeling the heat. Here at the transition between steppe grassland and taiga, plants and animals are confronted with a changing environment—and the outlook is not good for the herders who are crowding up from the south. Since the end of communism early in the last decade, the steppe has seen an explosion in livestock numbers—expanding up to threefold over the past 2 decades alone. The grasslands are on the verge of ecological collapse, says

Goulden. "The environmental problems are closing in on two fronts at once."

Tumorsuk's breath clouds his broad, ruddy face. He orders the scientists to gather stones. Then he pulls an ax from under his seat and disappears into the woods.

A blue pearl on ice

Sidelined by a chronic back injury, Goulden watches as his colleagues pile rocks next to the jeep, following Tumorsuk's mysterious order. If he's worried, he doesn't show it. Then again, Goulden, 70, is no stranger to hardship in the most sparsely populated country on Earth. Since helping launch long-term ecological studies of the Lake Hovsgol region in 1994, Goulden has dealt with many hassles. In 2006, fierce winds sank a motorboat that shuttled his team 100 kilometers to the northern shore. Fortunately, no one was injured. Sans boat, they now make a bone-rattling drive—8 hours in ideal conditions, 4 days at worst—to get supplies and scientists to field sites.

Hovsgol's forbidding location makes it an ecological wonderland. Of the world's 17 ancient lakes—those with the geologic fortune of existing continuously for millions of years—Hovsgol is the most pristine outside of Antarctica. The "blue pearl," as Mongolians call it, is untouched by the pollution that has spoiled most Eurasian lakes, and arid Mongolia wants to keep it that way.

Tumorsuk—who, like many Mongolians, does not use a last name—is one of 14 rangers charged with patrolling the Connecticut-sized nature reserve encompassing the lake. One of his main jobs is to track the hooves arriving from the south. Nomadic herders have plied the steppe since antiquity, grazing horses, yaks, goats, and sheep for a few months in one valley before

pulling up the poles of their felt-lined gers and moving on to the next. But the steppe has never supported the present horde: 35 million head of livestock, more than 10 times the human population of Mongolia. Tumorsuk must constantly cajole the herders not to overgraze the steppe around the lake.

While people around the world are giving up agrarianism for the city life, Mongolians are streaming in the opposite direction. During 8 decades of communism, a nation of herders was forcibly modernized. The results are a mixed bag: Although Mongolians enjoy one of the highest literacy rates in the world, their shamanistic traditions were all but exterminated, and the cities swelled, particularly Ulaanbaatar, where a third of Mongolians now live. But when the Soviet Union crumbled in 1991, newly democratic Mongolia suddenly lost its economic raison d'être. Bustling trade routes to Russia ebbed. Unemployment skyrocketed in the mid-1990s. Relations with China remain cautious.

"They fell back on what they know," says Goulden. "The problem is that most of the new herders don't have the know-how." The flow of traditional knowledge from one generation to the next—how to avoid conflicts with other herders, keep animals healthy, and avoid overgrazing, for instance—was severed. The limits of sustainability are being tested on a grand scale.

If land-use patterns were the only change, Mongolia's predicament would not be so dire. But now the land itself is changing. Winter temperatures in Mongolia have increased a staggering 3.6°C on average during the past 60 years. "The mountains are losing their snowcaps, and the glaciers on the northern shore are shrinking," says Goulden.

CREDIT: J. BOHANNON

This is no boon for agriculture. Although global warming has shifted the start of the growing season from June to May, precipitation is more erratic, Goulden says. Four of the worst drought years on record in Mongolia occurred in the past decade. And during the same period, intense storms have grown more frequent, according to a 2005 report by the United Nations Environment Programme. Flash floods erode the overgrazed steppe topsoil. But the worst weather condition is the dreaded dzud, an ice crust that forms over vegetation when rain freezes or melted snow refreezes. From 1999 to 2002, dzuds were a death sentence for 10 million animals that were unable to forage. The livestock losses spurred a wave of suicides among herders.

Where theory meets dirt

Dusk is settling on the taiga. Tumorsuk returns to the jeep, huffing as he lugs a 3-meter-long section of tree trunk on his shoulder. He drops it next to the jeep and digs around the sunken wheels with a shovel, clearing out sloppy muck. As they stand by helplessly, the ecologists' fingers and toes grow numb.

This muck is global warming's battleground. Lake Hovsgol straddles the southern edge of the Northern Hemisphere's permafrost. That edge is receding. As permafrost retreats deeper or disappears, the ground becomes a giant sponge that wicks water away from plant roots. That sets big changes in motion topside. "Taiga and permafrost always go together," says geophysicist Vladimir Romanovsky of the University of Alaska, Fairbanks. "You can't have one without the other." Hovsgol's taiga forest is growing patchier. And without the insulating tree cover, he says, soil warming accelerates.

Also worrying is a flash point created by drying soil and dying vegetation. Fires are a natural feature here, as shown by periodic dark bands in tree rings. But fires are growing more frequent and fierce, says Boldgiv. The worst-case scenario is that drought and wildfires converge in a regional conflagration. Huge swaths of taiga forest and steppe grasslands could be lost in a single summer, he says. There is no fire brigade out here.

Because of its likely impact on animal populations, a wholesale transition to semidesert would be "disastrous."
—CLYDE GOULDEN, ACADEMY OF NATURAL SCIENCES

Just what ecosystem might emerge from that apocalyptic scenario is a central question of the Hovsgol project. After a decade of research, ecologists have bad news and worse news. The bad news is that receding taiga and overgrazed steppe tend to leave shrubby wasteland in their wake. Islands of this "semidesert" of sparse plants and few grasses are expanding. Goulden is worried that this may warn of a wholesale transition to semidesert, which would be "disastrous," he says, because it supports a fraction of the animal density that grassland supports. And it could ruin the country's best source of drinking water if topsoil eroding into Hovsgol's tributaries spurs algal growth in the lake.

The worse news: This transformation could be a one-way ticket. A long-standing question in ecology is whether communities of species can be tipped into "alternative stable states." The steppe grasslands, for example, have proved for millennia to be a robust solution to life in cold, dry Mongolia. But once widespread conversion to semidesert occurs, it might be virtually impossible to reverse, says Goulden. In the taiga, even a temporary loss of permafrost, combined with extreme drought and fires, might be a point of no return, he says. The theory of alternative stable states is a mainstay of modern ecology, says ecologist Peter Petraitis of the University of Pennsylvania. But despite decades of experiments, "it remains just that—a theory." What is needed is the intense study of a real-world system, he says.

The American-Mongolian collaboration just received a boost to do that. Starting this year, the U.S. National Science Foundation is funding their work in Mongolia with a 5-year, $2.5 million grant, jointly led by Petraitis, Goulden, and Boldgiv. One task will be to widen the net of environmental monitoring they have established by compiling a map of permafrost, stream hydrology, and plant species distribution. To plug hard data into their ecological models of global warming, they will build chambers over experimental plots and measure the effects of temperature, plant cover, and other parameters on soil moisture and respiration. Also built into the grant is a training exchange between Mongolia and the United States. "The Mongolians know their environment better than anyone," says Goulden. "This must be an equal partnership."

With lightning speed, Tumorsuk carves the end of the log to precisely cup the jeep's wheel axle. Then, using another log as a fulcrum and the three ecologists as counterweight, he constructs an Archimedean lever. One side of the jeep rises from the mud. Tumorsuk grabs stones and plunges his naked hands into the water-filled wheel holes. Repeating the process on the opposite side and driving forward, the jeep moves half a meter before sinking back in. After 2 hours of levering, the jeep finds purchase and the team drives away into the night. The environment has been bested, for now.

—JOHN BOHANNON
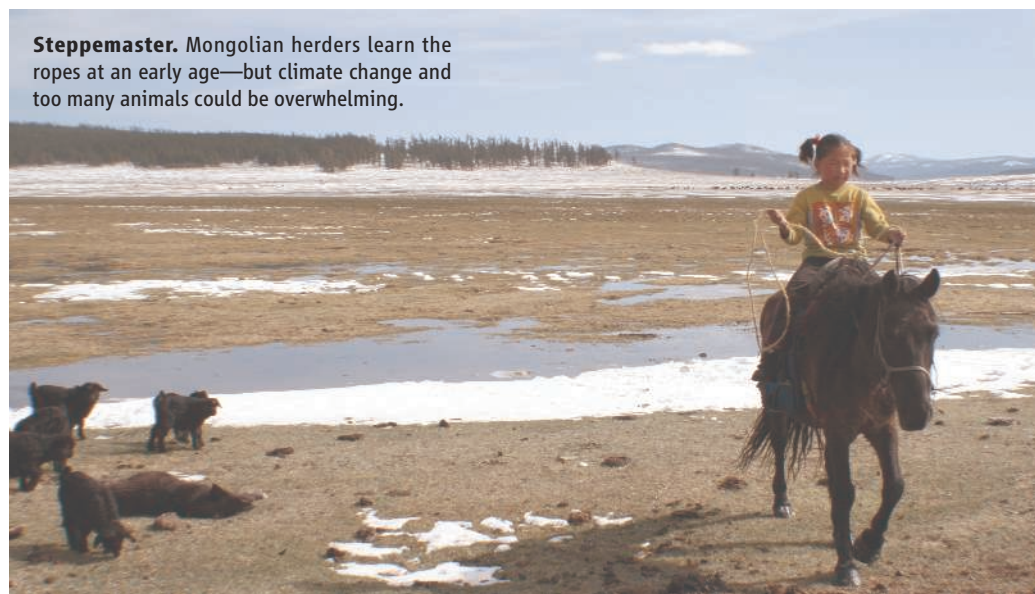
Steppemaster. Mongolian herders learn the ropes at an early age—but climate change and too many animals could be overwhelming.

CREDIT: J. BOHANNON

COMMENTARY

Phytoplankton history **571**	Challenges for water management **573**	Handling bones and bugs **576**

LETTERS I BOOKS I POLICY FORUM I EDUCATION FORUM I PERSPECTIVES

LETTERS

edited by Jennifer Sills

Retraction

WE WISH TO RETRACT OUR REPORT "COMPUTATIONAL DESIGN OF A BIOLOGICALLY ACTIVE enzyme" (*1*), which describes triose phosphate isomerase activity in a computationally redesigned ribose-binding protein (RBP) from *E. coli*. Dr. John P. Richard (Department of Chemistry, Department of Biochemistry, The State University of New York at Buffalo), to whom we provided clones encoding the novoTIM activity, has brought to our attention that the triose phosphate isomerase activity observed in our reported preparations can be attributed to a wild-type TIM impurity—seen in preparations that use a continuous rather than stepwise imidazole gradient (as in the original paper) or that add a second sepharose column. Richard's reanalysis has now also been confirmed by others in the Hellinga laboratory. The interpretations in the original report were based on lack of observed activity in mutant, engineered enzyme that bound substrate, but lacked catalytic residues. Variations in expression levels of designed proteins relative to the amount of contaminating endogenous protein might account for the pattern of observed activities that led to our erroneous conclusions. The in vivo experiments have not been reexamined.

We deeply regret that our report of a designed enzyme activity does not live up to closer scrutiny. Nevertheless, we remain optimistic that the problem of structure-based design of enzyme activity will be solved and that novel catalysts will be produced in conjunction with computationally based methods.

MARY A. DWYER,[1] LOREN L. LOOGER,[2] HOMME W. HELLINGA[3]

[1]Department of Pharmacology and Cancer Biology, Duke University Medical Center, Durham, NC 27710, USA. [2]Howard Hughes Medical Institute, Janelia Farm, Chevy Chase, MD 20815–6789, USA. [3]Department of Biochemistry, Duke University Medical Center, Durham, NC 27710, USA.

Reference
1. M. A. Dwyer, L. L. Looger, H. W. Hellinga, *Science* **304**, 1967 (2004).

Comparing Social Skills of Children and Apes

A RECENT RESEARCH ARTICLE BY E. HERRMANN *et al.* ("Humans have evolved specialized skills of social cognition: The cultural intelligence hypothesis," 7 September 2007, p. 1360) claims that compared with 2-year-old human children, great apes have equivalent technical skills but inferior social skills.

The study features an impressive battery of tests, seemingly administered in the same format to apes and children. However, when a human experimenter provides the social cues, the apes are at a disadvantage (*1*–*3*) because they are dealing with a species other than their own. This may not be as relevant for physical or technical problems, which focus on inanimate objects, but social tasks rely crucially on the relation between experimenter and subject. The reported findings are consistent with the idea that the methodology handicaps apes specifically in the social domain.

The differences between the setups for children and apes in this study appear multifold (*3*). Human children sit on or next to their parent (creating potential "Clever Hans" effects) and receive verbal instructions. They are used to dealing with strangers and are tested by a member of their own species. The apes are alone and confined, receive no verbal instructions, and are tested by a species not their own. We are not suggesting that human experimenters should never be used, but that the social skills that matter most for apes, especially with regards to social learning, are those shown with conspecific models.

In fact, evidence for ape-to-ape social learning is plentiful. Studies of wild chimpanzees in Africa have documented an impressive array of group-specific traditions attributed to social learning (*4*). Apes tested with a human model they have bonded with (*3*, *5*) or with a familiar member of their species (*6*–*9*) have demonstrated social learning that has extended to high-fidelity cultural transmission within and between groups. These findings conflict with the results as well as the central thesis of Herrmann *et al*.

We strongly urge testing of cognition in ecologically valid settings, such as testing social skills with conspecifics. The problem of the human model would be even more severe in relation to Herrmann *et al*.'s proposal to extend their test battery to more distantly related species.

FRANS B. M. DE WAAL,[1] CHRISTOPHE BOESCH,[2] VICTORIA HORNER,[1] ANDREW WHITEN,[3]

[1]Living Links, Yerkes National Primate Research Center, Emory University, Atlanta, GA 30322, USA. [2]Department of Primatology, Max Planck Institute for Evolutionary Anthropology, Deutscher Platz 6, 04 103, Leipzig, Germany. [3]Centre for Social Learning and Cognitive Evolution, School of Psychology, University of St. Andrews, St. Andrews, Fife KY16 9JP, UK.

References
1. F. B. M. de Waal, *Behav. Brain Sci.* **21**, 689 (1998).
2. F. de Waal, *The Ape and the Sushi Master: Cultural Reflections by a Primatologist* (Basic Books, New York, 2001).
3. C. Boesch, *J. Comp. Psychol.* **121**, 227 (2007).
4. A. Whiten *et al.*, *Nature* **399**, 682 (1999).
5. V. Horner, A. Whiten, *Anim. Cognit.* **8**, 164 (2005).
6. A. Whiten, V. Horner, F. B. M. de Waal, *Nature* **437**, 737 (2005).
7. V. Horner, A. Whiten, E. Flynn, F. B. M. de Waal, *Proc. Natl. Acad. Sci. U.S.A.* **103**, 13878 (2006).
8. K. E. Bonnie, V. Horner, A. Whiten, F. B. M. de Waal, *Proc. R. Soc. London Ser. B* **274**, 367 (2007).
9. A. Whiten *et al.*, *Curr. Biol.* **17**, 1038 (2007).

CREDIT: KAREN HUNT/CORBIS

Response

DE WAAL *ET AL.* QUESTION THE VALIDITY OF our finding that while chimpanzees, orangutans, and human children were equally skillful at cognitive tasks in the physical domain, human children were more skillful at cognitive tasks in the social domain. In their critique, there is a single hypothesis that could account for this discrepancy: The social tasks could be especially sensitive to the fact that only the children were tested by conspecifics. (Their other criticisms apply equally to tasks in the physical and social domains and so cannot account for the asymmetry.) We agree that studies with great apes should be conducted with apes interacting with apes whenever possible [e.g., (*1–3*)], but our Research Article (7 September 2007, p. 1360) is a broad assessment of the cognitive skills of three great ape species and so human experimenters, unfamiliar to all subjects, were necessary.

Recognizing this issue, we selected, whenever possible, cognitive tasks in the social domain that met two criteria: (i) There was previous research demonstrating no substantial difference when apes interacted with humans versus conspecifics; and (ii) apes had demonstrated some success in the past (see table S2 in the Supporting Online Material of our Research Article for the references). In addition, we measured the comfort level of each individual in our study when confronted with unfamiliar objects and humans. To our surprise, we found the human children to be more shy and less interested in interacting with unfamiliar human experimenters and objects than either of the ape species; moreover, within each species this assessment did not correlate with performance on the social tasks. Finally, in a recent survey by Boesch (*4*) of great ape cognitive research, it was concluded that across many studies "the use of a human experimenter did not seem to have an influence on the conclusion that humans perform better than chimpanzees."

Even though we reported on three scales of social cognition (all with similar results), de Waal *et al.* focus exclusively on social learning and the hypothesized effect of a human demonstrator on ape performance. We are great admirers of this team's studies of social learning using ape demonstrators, but the claim that human demonstrators are harmful for ape performance has no empirical basis: (i) This team's studies of ape social learning do not compare performance with a human and an ape demonstrator, and so do not address the issue (*5–9*); (ii) the one existing study that makes such a comparison finds no difference (*10*); (iii) the best evidence for action imitation, in a study by Whiten (*11*), used human demonstrators; and (iv) the best evidence that at least some apes reproduce actions in terms of their underlying intentions also comes from studies with human demonstrators (*12*). In our study, we do not claim that apes do not learn socially (in fact, they are skilled social learners); our finding is simply that human children prefer to follow the precise means of a demonstrated problem solution more often than do apes (who often prefer to solve it in their own way). This accords with the findings from almost all previous studies that include this comparison (*13–15*).

ESTHER HERRMANN,[1] JOSEP CALL,[1] MARÍA VICTORIA HERNÁNDEZ-LLOREDA,[2] BRIAN HARE,[1,3] MICHAEL TOMASELLO[1]

[1]Max Planck Institute for Evolutionary Anthropology, Leipzig D-04103, Germany. [2]Departamento de Metodología de las Ciencias del Comportamiento, Universidad Complutense de Madrid, Spain. [3]Department of Biological Anthropology and Anatomy, Duke University, Durham, NC 27705, USA.

References

1. B. Hare, J. Call, B. Agnetta, M. Tomasello, *Anim. Behav.* **59**, 771 (2000).
2. A. P. Melis, B. Hare, M. Tomasello, *Science* **311**, 1297 (2006).
3. K. Jensen, J. Call, M. Tomasello, *Science* **318**, 107 (2007).
4. C. Boesch, *J. Comp. Psychol.* **121**, 227 (2007).
5. V. Horner, A. Whiten, *Anim. Cognit.* **8**, 164 (2005).
6. A. Whiten, V. Horner, F. B. M. de Waal, *Nature* **437**, 737 (2005).
7. V. Horner, A. Whiten, E. Flynn, F. B. M. de Waal, *Proc. Natl. Acad. Sci. U.S.A.* **103**, 13878 (2006).
8. K. E. Bonnie, V. Horner, A. Whiten, F. B. M. de Waal, *Proc. R. Soc. London Ser. B* **274**, 367 (2007).
9. A. Whiten *et al.*, *Curr. Biol.* **17**, 1038 (2007).
10. J. Call, M. Tomasello, *J. Comp. Psychol.* **109**, 308 (1995).
11. D. M. Custance, A. Whiten, K.A. Bard, *Behaviour* **132**, 837 (1995).
12. M. Tomasello, M. Carpenter, *The Emergence of Social Cognition in Three Young Chimpanzees* (Monographs of the Society for Research in Child Development, 70, 2005).
13. A. Whiten, D. M. Custance, J.-C. Gomez, P. Teixidor, K. A. Bard, *J. Comp. Psychol.* **110**, 3 (1996).
14. K. Nagell, R. Olguin, M. Tomasello, *J. Comp. Psychol.* **107**, 174 (1993).
15. J. Call, M. Carpenter, M. Tomasello, *Anim. Cognit.* **8**, 151 (2005).

TECHNICAL COMMENT ABSTRACTS

COMMENT ON "Saturation of the Southern Ocean CO$_2$ Sink Due to Recent Climate Change"

Rachel M. Law, Richard J. Matear, Roger J. Francey

Unlike Le Quéré *et al.* (Reports, 22 June 2007, p. 1735), we do not find a saturating Southern Ocean carbon sink due to recent climate change. In our ocean model, observed wind forcing causes reduced carbon uptake, but heat and freshwater flux forcing cause increased uptake. Our inversions of atmospheric carbon dioxide show that the Southern Ocean sink trend is dependent on network choice.

Full text at www.sciencemag.org/cgi/content/full/319/5863/570a

COMMENT ON "Saturation of the Southern Ocean CO$_2$ Sink Due to Recent Climate Change"

Kirsten Zickfeld, John C. Fyfe, Michael Eby, Andrew J. Weaver

We disagree with the conclusion of Le Quéré *et al.* (Reports, 22 June 2007, p. 1735) that poleward intensifying winds could continue to weaken the Southern Ocean sink in the future. We argue that altered winds, along with rising atmospheric carbon dioxide, will likely increase the efficiency of this sink in the 21st century.

Full text at www.sciencemag.org/cgi/content/full/319/5863/570b

RESPONSE TO COMMENTS ON "Saturation of the Southern Ocean CO$_2$ Sink Due to Recent Climate Change"

Corinne Le Quéré, Christian Rödenbeck, Erik T. Buitenhuis, Thomas J. Conway, Ray Langenfelds, Antony Gomez, Casper Labuschagne, Michel Ramonet, Takakiyo Nakazawa, Nicolas Metzl, Nathan P. Gillett, Martin Heimann

We estimated a weakening of the Southern Ocean carbon dioxide (CO$_2$) sink since 1981 relative to the trend expected from the large increase in atmospheric CO$_2$. We agree with Law *et al.* that network choice increases the uncertainty of trend estimates but argue that their network of five locations is too small to be reliable. A future reversal of Southern Ocean CO$_2$ saturation as suggested by Zickfeld *et al.* is possible, but only at high atmospheric CO$_2$ concentrations, and the effect would be temporary.

Full text at www.sciencemag.org/cgi/content/full/319/5863/570c

Letters to the Editor

Letters (~300 words) discuss material published in *Science* in the previous 3 months or issues of general interest. They can be submitted through the Web (www.submit2science.org) or by regular mail (1200 New York Ave., NW, Washington, DC 20005, USA). Letters are not acknowledged upon receipt, nor are authors generally consulted before publication. Whether published in full or in part, letters are subject to editing for clarity and space.

Comment on "Saturation of the Southern Ocean CO₂ Sink Due to Recent Climate Change"

Rachel M. Law,[1]* Richard J. Matear,[2] Roger J. Francey[1]

Unlike Le Quéré et al. (Reports, 22 June 2007, p. 1735), we do not find a saturating Southern Ocean carbon sink due to recent climate change. In our ocean model, observed wind forcing causes reduced carbon uptake, but heat and freshwater flux forcing cause increased uptake. Our inversions of atmospheric carbon dioxide show that the Southern Ocean sink trend is dependent on network choice.

Le Quéré et al. (1) reported that the Southern Ocean sink of CO_2 has weakened since 1981 despite the increase in atmospheric CO_2 levels. To further test their assessment, we estimated the carbon flux from the Southern Ocean using a Bayesian synthesis inversion method (2, 3). The inversion gives very similar results to (1) for the trend and interannual variability (IAV) in the Southern Ocean CO_2 sink provided that the same network of atmospheric CO_2 data is used. However, we also found that the estimated trend is dependent on the network choice.

Our control inversion used nine data records (4, 5) from five locations (6) [a smaller network than (1)], and we compared this with inversions that added Amsterdam Island (AMS) and Ascension Island (ASC) data used in (1). Inversions of synthetic data (7) indicated that any of these cases should give estimates of Southern Ocean CO_2 flux IAV and trend that are consistent with the fluxes used to generate the synthetic data. Figure 1 shows the three estimates of the annual mean Southern Ocean carbon flux and the annual mean uncertainty for the control case. The interannual variations are similar between cases, but the trend in sink between 1981 and 2004 [the period used in (1)] is not. The control case gives an increasing CO_2 sink with a trend of -0.092 ± 0.084 petagrams of carbon (Pg C) year^{-1} decade^{-1} (8). Adding the AMS data to the inversion weakens the negative trend (-0.012 ± 0.081 Pg C year^{-1} decade^{-1}), whereas adding both AMS and ASC data gives a positive trend (0.049 ± 0.076 Pg C year^{-1} decade^{-1}) similar to (1).

The control inversion trend is consistent with that produced by an ocean carbon model (9) run

with constant 1948 wind, heat flux, and freshwater flux forcing but with increasing atmospheric CO_2 (Fig. 1). This produces a trend for the period 1981 to 2002 of -0.088 ± 0.003 Pg C year^{-1} decade^{-1}. Hence, our control inversion does not produce a trend that is significantly different from that expected due to increasing atmospheric CO_2. Our inversion case closest to (1) (with AMS and ASC) is significantly different from the ocean model trend at the 95% level. The significance is less than in (1) because we included the flux
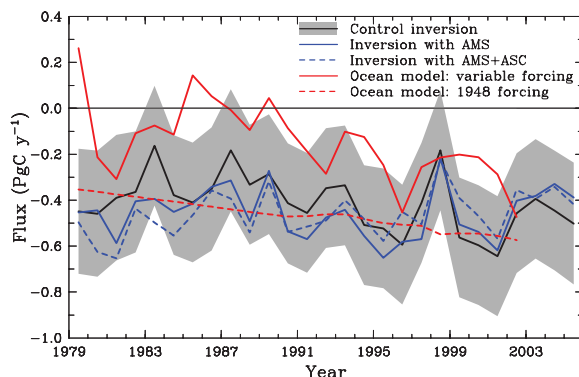
Fig. 1. Annual mean Southern Ocean CO_2 flux from inversions using the control network (black), adding AMS (solid blue), and adding AMS and ASC (dotted blue) and from an ocean model forced with observed winds and fluxes (solid red) and constant wind and fluxes (dotted red). The shaded region shows the ±1 SD uncertainty on the fluxes for the control inversion. The long-term mean flux cannot be reliably estimated from the inversions (7), so the mean offset between the ocean model with variable forcing and inversion fluxes is not considered significant.

uncertainty from the inversion in the calculation of the trend standard deviation. The inversion flux uncertainty (Fig. 1) is determined primarily from the data uncertainty used in the inversion [0.3 to 0.5 parts per million (ppm)] (10), which encompasses the ability to model CO_2 at the sites and measurement error (estimated to be 0.2 to 0.4 ppm based on differences between colocated CO_2 records at the South Pole and Samoa).

Figure 1 also shows the Southern Ocean carbon uptake simulated by the ocean carbon

model forced with observed winds (11) and heat and freshwater fluxes from 1948 to 2002. The increase in sink over the last two decades is slightly larger (-0.139 ± 0.043 Pg C year^{-1} decade^{-1}), but not significantly different from that seen in the constant forcing ocean run and in the control inversion. In the variable forcing case, we find compensating trends from the wind forcing and from the flux forcing; variable wind forcing gives the saturating Southern Ocean sink found by (1) in their ocean model simulation. In contrast to the findings of Le Quéré et al. (1), variable flux forcing increases CO_2 uptake in the Southern Ocean. Our fluxes (12) differ from (1), but it is difficult at present to determine which flux fields are more reliable. Clearly, the ocean model simulations are highly sensitive to the choice of flux fields and how they are used in the model.

The ocean model results suggest that the control inversion trend may be more realistic than the trend from the inversions including AMS and ASC. This is supported by inspection of the CO_2 records for ASC (13) and AMS (14). High proportions of positive outliers in ASC samples before 1991 result in poor definition of the seasonal cycle and increased uncertainty in annual averages. Also, differences from South Pole (ASC-SPO) compare well with Samoa–South Pole differences (SMO-SPO) after 1989, but from 1982 to 1986 the ASC data appear 0.7 ± 0.4 ppm higher than expected if SMO and ASC are responding to the same long-term trends. At AMS, after 1999, springtime values tend to be lower than those from other Southern Hemisphere sites by up to 0.5 ppm, with the August-September 2004 data appearing low by an additional 0.7 ppm.

To assess the impact of these apparent anomalies, we performed an inversion with ASC data reduced by 0.5 ppm from 1981 to 1986, and AMS data increased by 0.2 ppm from 1999 to 2005. The Southern Ocean source increased by about 0.12 Pg C year^{-1} from 1981 to 1986 and decreased by about 0.08 Pg C year^{-1} from 1999 to 2005. These differences are smaller than the source uncertainty from the inversion (0.23 Pg C year^{-1}) but change the 1981 to 2004 sink trend from positive to negative (-0.037 Pg C year^{-1} decade^{-1}). Unless we can be confident that the changes in CO_2 at AMS and ASC relative to SPO are driven by changes in surface fluxes rather than measurement or sampling errors, we cannot estimate a robust trend in the Southern Ocean sink.

Both our inversion of atmospheric CO_2 and our ocean model indicate that the Southern Ocean trend found by (1) is not robust and that there is insufficient evidence to conclude that the Southern Ocean sink of CO_2 has saturated as a result of recent climate change. The inversion of atmospheric CO_2 remains a vital method for monitoring the response of the natural sources and sinks

[1]Wealth from Oceans Flagship, Commonwealth Scientific and Industrial Research Organisation (CSIRO), PMB 1, Aspendale, Victoria 3195, Australia. [2]Wealth from Oceans Flagship, CSIRO, GPO Box 1538, Hobart, Tasmania 7001, Australia.

*To whom correspondence should be addressed. E-mail: rachel.law@csiro.au

of CO_2 to anthropogenic emissions and climate change, but the ability to detect long-term trends requires careful use of the atmospheric measurements and greater resources to provide a denser network in space and time. Our ocean model simulations show that changes in heat fluxes, freshwater fluxes, and winds all substantially affect the trend in Southern Ocean CO_2 uptake, and how these fields are used in the ocean model is important. Obtaining consistent interannual variations in ocean fluxes from the ocean model and atmospheric inversions would increase confidence that model processes are well represented.

References and Notes

1. C. Le Quéré *et al.*, *Science* **316**, 1735 (2007).
2. D. F. Baker *et al.*, *Global Biogeochem. Cycles* **20**, GB1002 (2006).
3. We solve for 116 regions globally, 6 of which make up the Southern Ocean region, using monthly mean CO_2 mixing ratio measurements.
4. GLOBALVIEW-CO_2, Cooperative Atmospheric Data Integration Project–Carbon Dioxide, CD-ROM, NOAA GMD, Boulder, CO (2006). [Also available on Internet via anonymous FTP to ftp.cmdl.noaa.gov, Path: ccg/co2/ GLOBALVIEW.]
5. C. D. Keeling, T. P. Whorf, *Trends: A Compendium of Data on Global Change* (Carbon Dioxide Information Analysis Center, Oak Ridge National Laboratory, U.S. Department of Energy, Oak Ridge, TN, 2005).
6. Control inversion sites with data type and uncertainty used in the inversion: Barrow (71°N, 157°W), NOAA (National Oceanic and Atmospheric Administration) in situ, 0.7 ppm; Mauna Loa (20°N, 156°W), NOAA in situ, 0.5 ppm; Samoa (14°S, 171°W), NOAA flask and in situ, SIO (Scripps Institution of Oceanography) flask, 0.5 ppm; Palmer Station (65°S, 64°W), NOAA flask, 0.3 ppm; South Pole (90°S, 25°W), NOAA flask and in situ, SIO flask, 0.3 ppm. Additional sites for sensitivity tests: Amsterdam Island (38°S, 78°E), LSCE (Laboratoire des Sciences du Climat et de l'Environnement) in situ, 0.3 ppm; Ascension Island (8°S, 14°W), NOAA flask, 0.5 ppm. The NOAA and LSCE records were from GLOBALVIEW-CO_2, and four pseudo-weekly values were averaged to give monthly means. Gap-filled values were used, but the data uncertainty was increased by up to a factor of five when this occurred.
7. The inversion set-up was tested using a synthetic data test for 1971 to 1999 in which interannually varying CO_2 fluxes were input into an atmospheric model to create CO_2 concentration time series. The fluxes were retrieved using an inversion performed with a different atmospheric model. The forward simulation included interannually varying meteorology, whereas the inversion did not. We found that the control inversion was unable to retrieve the long-term mean flux but was successful in retrieving the interannual variations (with a slight underestimate in variability) and sink trend (−0.063 Pg C year^{-1} decade^{-1} for 1976 to 1999 compared with −0.068 Pg C year^{-1} decade^{-1} for the input fluxes). Inversions including Amsterdam Island without or with Ascension Island gave comparable results with slightly smaller trends (−0.049 and −0.050 Pg C year^{-1} decade^{-1}, respectively).
8. The standard deviation of the trend incorporates two components that we consider to be independent: interannual variability in the estimated annual mean sources and the uncertainty estimated by the inversion for each of those annual means. The variance of the trend is the sum of the variance of the annual mean sources around the trend line and the variance from the inversion estimate divided by the sum of the squared deviations of the years from the mean year (1993). The calculation assumes that the annual mean sources are normally distributed around the trend line. For the ocean model simulations, only the interannual variability component is used.
9. A. Lenton, R. J. Matear, *Global Biogeochem. Cycles* **21**, GB2016 (2007).
10. The number of sites constraining the Southern Ocean region also determines the flux uncertainty. For example, a network with five extra sites in the Southern Ocean and Antarctic region (such as currently available) would reduce the uncertainty by about 25%.
11. E. Kalnay *et al.*, *Bull. Am. Meteorol. Soc.* **77**, 437 (1996).
12. Our simulation uses heat and freshwater fluxes from the National Centers for Environmental Prediction reanalysis, with additional fluxes due to restoring the simulated sea surface temperature and salinity fields to observations on a 30-day time scale.
13. P. P. Tans, T. J. Conway, *World Data Centre for Greenhouse Gases* (Japan Meteorological Agency, Tokyo, 2007); http://gaw.kishou.go.jp/wdcgg.html.
14. M. Ramonet, M. Schmidt, P. Ciais, V. Kazan, S. G. Jennings, *World Data Centre for Greenhouse Gases* (Japan Meteorological Agency, Tokyo, 2007); http://gaw.kishou.go.jp/wdcgg.html.

BIOLOGICAL OCEANOGRAPHY

Drifters Through Time

Robert Riding

In hindsight, it is not surprising that floating cells evolved to take advantage of sunlight, water, and carbon dioxide at the surface of a habitable planet such as Earth. Phytoplankton—myriad small, mostly unicellular algae and bacteria that occupy the surface waters of seas and lakes—dominate aquatic primary production, and over time they have radically altered Earth's atmosphere. But reconstructing their multimillion-year history is a daunting task. Those with tough shells or cysts have a good fossil record, whereas others have left scarcely a trace. In *Evolution of Primary Producers in the Sea*, editors Paul Falkowski and Andrew Knoll take a forthright approach to this challenge. They rely on resourceful detective work, integrate all available strands of biological and geological information, and wherever possible emphasize large-scale interconnections between life and environment.

The star of the show is oxygenic photosynthesis—the physiological process employed by most present-day phytoplankton—which uses sunlight as energy to create organic matter from carbon dioxide and water, releasing oxygen as a by-product. As a result, atmospheric carbon dioxide is buried as organic carbon, while oxygen is liberated to permeate the oceans and atmosphere. Phytoplankton have been no more immune to the dramatic consequences of these effects than other organisms, and how they responded to the changes they helped create is one of the volume's fascinating central themes.

The authors work hard to place phytoplankton evolution in its geological perspective, but the fossil record can be inscrutable. Some key questions, such as when cyanobacteria (the organisms believed to have originated oxygenic photosynthesis) first appeared, are surprisingly hard to resolve. We thought we had the answer until serious doubts were raised about the nature of 3500-million-year-old cyanobacteria-like microfossils in Australia (*1*). Currently it seems more likely that cyanobacteria and oxygenic photosynthesis arose later, perhaps ~2900 million years ago (Ma) (*2*).

The inception of the phytoplankton record, ~1700 Ma, raises further questions by introducing enigmatic organic walled microfossils that look algal—some resemble present-day dinoflagellates and others prasinophyte green algae—but whose precise affinities are unknown. These fossils guard their secrets so well that they are called acritarchs: the only major group of organisms defined as being "of uncertain origin." Acritarchs dominated the phytoplankton record until 250 Ma, so the uncertainties that surround them hamper understanding of an immensely long period in the evolution of marine life.

Even so, large-scale patterns are discernible in acritarch history: very slow diversification
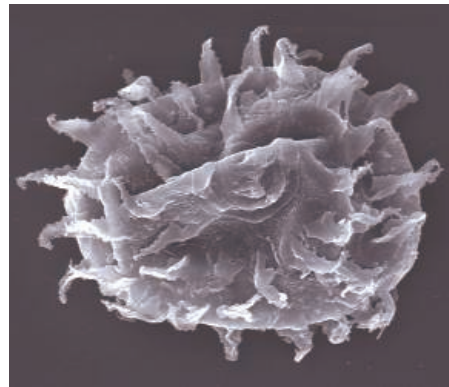
Origin uncertain. The acritarch *Peteinosphaeridium septuosum* from the Sylvan Shale (Late Ordovician, Oklahoma).

prior to 550 Ma, acme at 450 Ma, followed by a decline that was steepest 360 Ma. The protracted initial diversification from 1700 Ma to 550 Ma might reflect the gradual pace of Earth's oxygenation. Persistent deep ocean anoxia may have retarded nutrient release until oxygen levels finally rose ~550 Ma, stimulating algal phytoplankton and animal diversification as suspension feeders and zooplankton grazers evolved.

How eukaryotes acquired oxygenic photosynthesis after it was developed by cyanobacteria is almost the stuff of science fiction. Nonphotosynthetic organisms reinvented themselves as red and green algae by engulfing cyanobacterial cells and transforming them into photosynthetic organelles, plastids. The process was repeated when red and green algae were in turn engulfed as secondhand plastids. Red-algal plastids, for example, were incorporated into coccolithophores and diatoms. Dinoflagellates have the distinction of receiving plastids thirdhand, from diatoms and other algae. This remarkable "pass the plastid" history, and much else that is known about phytoplankton, is deduced from present-day organisms, but fossil evidence is required to confirm absolute ages and reconstruct past ecologic interactions. During the 300 million years of the Paleozoic era, ocean oxygenation continued to shift the availability of trace metals essential for cell biochemistry in a direction that favored the red-algal plastids acquired by dinoflagellates, coccolithophorids, and diatoms in the Mesozoic era. This suggests an intimate link between seawater chemistry and the rise of these important groups, as well as the continuing overriding influence of oxygen on algal evolution.

After the difficulties posed by acritarchs, there is an almost-palpable sigh of relief when, at 250 Ma, the story finally reaches the familiar Mesozoic algae, whose resistant cysts and shells permit their record to be tracked in detail to the present day. But an important group is missing. Cyanobacteria outnumber if not outweigh present-day algal phytoplankton in abundance, but they are unknown as body fossils. Many of their cells are less than two micrometers in size. Even in present-day seas, the importance of such picophytoplankton went unrecognized until 30 years ago (*3*). Reconstructing the history of cyanobacterial phytoplankton is a major challenge.

The volume rightly emphasizes the effects of Earth's oxygenation on phytoplankton evolution. After all, these organisms were presumably largely responsible for the significant oxygen increase that occurred before terrestrial plants greened the continents. At the same time, decrease in carbon dioxide has also strongly influenced phytoplankton, because inorganic carbon is fundamental for photosynthesis. The effects of low carbon dioxide levels is one of the few topics that I would have given more prominence in the volume. After all, many phytoplankton have mechanisms to deal with present-day low levels of carbon dioxide, and this ability is likely to have played an important role in phytoplankton history. Cyanobacteria, for example, overcome carbon limitation by pumping bicarbonate into their cells, and they may have been doing this since carbon dioxide declined substantially ~360 Ma (*4*).

Any attempt to tell the phytoplankton story

The reviewer is at the School of Earth, Ocean and Planetary Sciences, Cardiff University, Cardiff CF10 3YE, UK. E-mail: riding@cardiff.ac.uk

CREDIT: COURTESY OF REED WICANDER/CENTRAL MICHIGAN UNIVERSITY AND GEOFFREY PLAYFORD/THE UNIVERSITY OF QUEENSLAND

is welcome. The authors of *Evolution of Primary Producers in the Sea* go further and succeed in making a complex and at times perplexing subject accessible and exciting. Comprehensive and authoritative, their explorations of fundamental questions and global geobiological trends are engaging and thought provoking. The volume will be influential, and it should signal a turning point in phytoplankton research.

References
1. M. D. Brasier *et al.*, *Nature* **416**, 76 (2002).
2. T. Cavalier-Smith, *Philos. Trans. R. Soc. London Ser. B* **361**, 969 (2006).
3. J. B. Waterbury *et al.*, *Nature* **277**, 293 (1979).
4. M. R. Badger, G. D. Price, *J. Exp. Bot.* **54**, 609 (2003).

ECOLOGY

Give Way to the Migrants

Thomas Alerstam

Migration represents a spectacularly successful strategy among animals, providing access to a richness of ephemeral and seasonal resources that can sustain large populations. Its importance for promoting abundance was stressed by Alfred Russel Wallace in his 1858 paper that set forth the fundamentals of natural selection in biological evolution and stirred Darwin to finally publish his long-considered ideas (*1*). Wallace pointed to the example of the passenger pigeon, which—in spite of its limited fecundity and flagrant exposure to predation—reached its immense abundance through rapid long-distance movements from depleted to fresh feeding grounds. The example illustrates, he argued, that animal populations "can never increase beyond the supply of food in the least favourable season." What he did not realize at that time was the passenger pigeon's great vulnerability to human exploitation. Within Wallace's lifetime (1823–1913), the species plummeted from tens of millions of birds. The last-known individual died in captivity in 1914.

The view of animal migration as a phenomenon of abundance and vulnerability forms the central theme of David Wilcove's *No Way Home*. His alarming message is that

The reviewer is at the Department of Animal Ecology, Lund University, Ecology Building, Lund, SE-22362, Sweden. E-mail: thomas.alerstam@zooekol.lu.se

around the world great animal migrations are disappearing. Thus, international conservation efforts are urgently needed to save the migrants from the devastating effects of overexploitation, habitat destruction, human-created obstacles, and climate change.

Animals traveling thousands or tens of thousands of kilometers in the air, on land, or in water inspire much awe. To complete its annual return journey between northerly breeding latitudes and tropical winter regions, a tiny songbird must keep to seasonal and daily timetables, change its physiological machinery between phases of fuel consumption and fuel deposition, vary flight steps and fuel loads in relation to the crossing of benign or hostile regions, find its way by compass and navigation systems, negotiate weather and winds, and correctly adjust flight speed and altitude. The bird's endowment with all necessary instructions represents a striking manifestation of the accomplishments achieved by biological evolution.

Wilcove, an ecologist at Princeton University, presents elegant and informed accounts of migrations in various taxa: birds (the New and Old World systems of billions of songbirds traveling to and from tropical winter quarters, red knots flying between the latitudinal extremes of the American continents, and bellbirds moving down and up the slopes in Central American cloud forests), insects (dragonflies that behave like migrating birds; monarch butterflies that depart each spring from high-altitude fir forests in Michoacán, Mexico, to start a multigenerational annual cycle of movement across North America; and now-extinct Rocky Mountain locusts that once moved in swarms of millions), terrestrial mammals (wildebeest of the Serengeti, springbok of South Africa, white-eared kob of Sudan, and bison and pronghorn of North America), sea mammals (right whale in the Atlantic and gray whale in the Pacific), sea turtles, and fish (Atlantic and Pacific salmon).

For each case, Wilcove takes us into the field to meet the animals (or to the scene of now-extinct migrations), often in company with researchers conducting exciting projects. Migration studies are currently in a phase of dynamic development, with novel tracking, physiological, and molecular techniques (*2*). In addition, the author provides fascinating stories of the animals' natural history, glimpses of recent scientific discoveries about migration performance and navigation mechanisms, and his-

torical sketches. He also describes population trends and describes the threats and conservation efforts. These strands are skillfully woven together, making his comprehensive perspective on animal travelers a delight to read.

Some of the migrants' predicaments stem from the complexity in seasonally and spatially shifting uses of resources. Increased specialization often goes hand in hand with increased vulnerability. However, the picture is not altogether dark. Some migratory populations, such as the gray whale, have shown encouraging recoveries. In recent decades, reduced persecution and the banning of toxins have led to the comeback of many birds of prey, including both short- and long-distance migrants. Changing their migration routes to exploit new resources provided by farming, some populations of geese and cranes have dramatically expanded. Their opportunistic flexibility is facilitated by learning; knowledge of migration routes is transferred between generations that travel together in families or mixed flocks (*3*). For still other species, migration may promote range expansion, leading to the establishment of new travel routes and the colonization of new breeding destinations. I would have appreciated more discussion of factors that differ between declining and expanding migratory populations. How important in this respect are cultural versus genetic evolution of migratory routes, short versus long migration distances, and levels of complexity in the annual cycle and habitat requirements?

Absorbing and thought provoking, *No Way Home* deserves to be widely read and used to promote conservation action. It illustrates the importance of science for deepening our appreciation of animal migrations and for guiding our efforts to preserve them. There is no conflict between scientific exploration of migratory mechanisms and connectivity and aesthetic marveling at the superb arrangements of nature. The investigation of animal migration is a major challenge in biology, more fascinating and urgent than ever. Wilcove urges us to proactively protect threatened migration systems while the migrants are still abundant.

References
1. A. R. Wallace, *Proc. Linn. Soc. London* **3**, 53 (1858).
2. M. S. Webster, P. P. Marra, S. M. Haig, S. Bensch, R. T. Holmes, *Trends Ecol. Evol.* **17**, 76 (2002).
3. W. J. Sutherland, *J. Avian Biol.* **29**, 441 (1998).

CLIMATE CHANGE

Stationarity Is Dead: Whither Water Management?

Climate change undermines a basic assumption that historically has facilitated management of water supplies, demands, and risks.

P. C. D. Milly,[1]* Julio Betancourt,[2] Malin Falkenmark,[3] Robert M. Hirsch,[4] Zbigniew W. Kundzewicz,[5] Dennis P. Lettenmaier,[6] Ronald J. Stouffer[7]

Systems for management of water throughout the developed world have been designed and operated under the assumption of stationarity. Stationarity—the idea that natural systems fluctuate within an unchanging envelope of variability—is a foundational concept that permeates training and practice in water-resource engineering. It implies that any variable (e.g., annual stream-flow or annual flood peak) has a time-invari-ant (or 1-year–periodic) probability density function (pdf), whose properties can be esti-mated from the instrument record. Under sta-tionarity, pdf estimation errors are acknowl-edged, but have been assumed to be reducible by additional observations, more efficient estimators, or regional or paleohydrologic data. The pdfs, in turn, are used to evaluate and manage risks to water supplies, water-works, and floodplains; annual global invest-ment in water infrastructure exceeds U.S.\$500 billion (1).

The stationarity assumption has long been compromised by human disturbances in river basins. Flood risk, water supply, and water quality are affected by water infra-structure, channel modifications, drainage works, and land-cover and land-use change. Two other (sometimes indistinguishable) challenges to stationarity have been exter-nally forced, natural climate changes and low-frequency, internal variability (e.g., the Atlantic multidecadal oscillation) enhanced by the slow dynamics of the oceans and ice sheets (2, 3). Planners have tools to adjust their analyses for known human distur-bances within river basins, and justifiably or not, they generally have considered natural change and variability to be sufficiently small to allow stationarity-based design.

[1]U.S. Geological Survey (USGS), c/o National Oceanic and Atmospheric Administration (NOAA) Geophysical Fluid Dynamics Laboratory, Princeton, NJ 08540, USA. [2]USGS, Tucson, AZ 85745, USA. [3]Stockholm International Water Institute, SE 11151 Stockholm, Sweden. [4]USGS, Reston, VA 20192, USA. [5]Research Centre for Agriculture and Forest Environment, Polish Academy of Sciences, Poznań, Poland, and Potsdam Institute for Climate Impact Research, Potsdam, Germany. [6]University of Washington, Seattle, WA 98195, USA. [7]NOAA Geophysical Fluid Dynamics Laboratory, Princeton, NJ 08540, USA.

*Author for correspondence. E-mail: cmilly@usgs.gov.

An uncertain future challenges water planners.

In view of the magnitude and ubiquity of the hydroclimatic change apparently now under way, however, we assert that stationarity is dead and should no longer serve as a central, default assumption in water-resource risk assessment and planning. Finding a suitable successor is crucial for human adaptation to changing climate.

How did stationarity die? Stationarity is dead because substantial anthropogenic change of Earth's climate is altering the means and extremes of precipitation, evapo-transpiration, and rates of discharge of rivers (4, 5) (see figure, above). Warming aug-ments atmospheric humidity and water transport. This increases precipitation, and possibly flood risk, where prevailing atmo-spheric water-vapor fluxes converge (6). Rising sea level induces gradually height-ened risk of contamination of coastal fresh-water supplies. Glacial meltwater temporar-ily enhances water availability, but glacier and snow-pack losses diminish natural sea-sonal and interannual storage (7).

Anthropogenic climate warming appears to be driving a poleward expansion of the subtropical dry zone (8), thereby reducing runoff in some regions. Together, circulatory and thermodynamic responses largely explain the picture of regional gainers and losers of sustainable freshwater availability

that has emerged from climate models (see figure, p. 574).

Why now? That anthropogenic climate change affects the water cycle (9) and water supply (10) is not a new finding. Nevertheless, sensible objections to discarding stationarity have been raised. For a time, hydroclimate had not demonstrably exited the envelope of natu-ral variability and/or the effective range of optimally operated infrastructure (11, 12). Accounting for the substantial uncertainties of climatic parameters estimated from short records (13) effectively hedged against small climate changes. Additionally, climate projec-tions were not considered credible (12, 14).

Recent developments have led us to the opinion that the time has come to move beyond the wait-and-see approach. Pro-jections of runoff changes are bolstered by the recently demonstrated retrodictive skill of cli-mate models. The global pattern of observed annual streamflow trends is unlikely to have arisen from unforced variability and is consis-tent with modeled response to climate forcing (15). Paleohydrologic studies suggest that small changes in mean climate might produce large changes in extremes (16), although attempts to detect a recent change in global flood frequency have been equivocal (17, 18). Projected changes in runoff during the multidecade lifetime of major water infra-structure projects begun now are large enough to push hydroclimate beyond the range of historical behaviors (19). Some regions have little infrastructure to buffer the impacts of change.

Stationarity cannot be revived. Even with aggressive mitigation, continued warming is very likely, given the residence time of atmospheric CO_2 and the thermal inertia of the Earth system (4, 20).

A successor. We need to find ways to identify nonstationary probabilistic models of relevant environmental variables and to use those models to optimize water systems. The challenge is daunting. Patterns of change are complex; uncertainties are large; and the knowledge base changes rapidly.

Under the rational planning framework advanced by the Harvard Water Program (21, 22), the assumption of stationarity was

combined with operations research, statistics, and welfare economics to formulate design problems as trade-offs of costs, risks, and benefits dependent on variables such as reservoir volume. These trade-offs were evaluated by optimizations or simulations using either long historical streamflow time series or stochastic simulations of streamflow based on properties of the historical time series.

This framework can be adapted to changing climate. Nonstationary hydrologic variables can be modeled stochastically to describe the temporal evolution of their

Human influences. Dramatic changes in runoff volume from ice-free land are projected in many parts of the world by the middle of the 21st century (relative to historical conditions from the 1900 to 1970 period). Color denotes percentage change (median value from 12 climate models). Where a country or smaller political unit is colored, 8 or more of 12 models agreed on the direction (increase versus decrease) of runoff change under the Intergovernmental Panel on Climate Change's "SRES A1B" emissions scenario.

pdfs, with estimates of uncertainty. Methods for estimating model parameters can be developed to combine historical and paleo-hydrologic measurements with projections of multiple climate models, driven by multiple climate-forcing scenarios.

Rapid flow of such climate-change information from the scientific realm to water managers will be critical for planning, because the information base is likely to change rapidly as climate science advances during the coming decades. Optimal use of available climate information will require extensive training of (both current and future) hydrologists, engineers, and managers in nonstationarity and uncertainty. Reinvigorated development of methodology may require focused, interdisciplinary efforts in the spirit of the Harvard Water Program.

A stable institutional platform for climate predictions and climate-information delivery may help (23). Higher-resolution simulations of the physics of the global land-atmosphere system that focus on the next 25 to 50 years are crucial. Water managers who are developing plans for their local communities to adapt to climate change will not be best served by a model whose horizontal grid has divisions measured in hundreds of kilometers. To facilitate information transfer in both directions between climate science and water management, the climate models need to include more explicit and faithful representation of surface- and ground-water processes, water infrastructure, and water users, including the agricultural and energy sectors.

Treatments of land-cover change and land-use management should be routinely included in climate models. Virtual construction of dams, irrigation of crops, and harvesting of forests within the framework of climate models can be explored in a collaboration between climate scientists and resource scientists and managers.

Modeling should be used to synthesize observations; it can never replace them. Assuming climatic stationarity, hydrologists have periodically relocated stream gages (24) so that they could acquire more perspectives on what was thought to be a fairly constant picture. In a nonstationary world, continuity of observations is critical.

The world today faces the enormous, dual challenges of renewing its decaying water infrastructure (25) and building new water infrastructure (26). Now is an opportune moment to update the analytic strategies used for planning such grand investments under an uncertain and changing climate.

References and Notes

1. R. Ashley, A. Cashman, in *Infrastructure to 2030: Telecom, Land Transport, Water and Electricity* (Organization for Economic Cooperation and Development, Paris, 2006).
2. R. H. Webb, J. L. Betancourt, *U.S. Geol. Surv. Water-Supply Paper* **2379**, 1 (1992).
3. C. A. Woodhouse, S. T. Gray, D. M. Meko, *Water Resour. Res.* **42**, W05415 (2006).
4. Intergovernmental Panel on Climate Change (IPCC), in *Climate Change 2007: The Physical Science Basis, Contribution of Working Group (WG) 1 to the Fourth Assessment Report of the IPCC (AR4)*, S. Solomon et al., Eds. (Cambridge Univ. Press, New York, 2007), pp. 1–18; www.ipcc.ch/press/index.htm.
5. IPCC, in *Climate Change 2007: Climate Change Impacts, Adaptation and Vulnerability, Contribution of WG2 to AR4*, M. L. Parry et al., Eds. (Cambridge Univ. Press, New York, 2007), pp. 1–16.
6. I. M. Held, B. J. Soden, *J. Clim.* **19**, 5686 (2006).
7. T. P. Barnett, J. C. Adam, D. P. Lettenmaier, *Nature* **438**, 303 (2005).
8. J. Lu, G. A. Vecchi, T. Reichler, *Geophys. Res. Lett.* **34**, L06805 (2007).
9. S. Manabe, R. J. Stouffer, *J. Geophys. Res.* **85**, 5529 (1980).
10. P. S. Eagleson, in *Scientific Basis of Water-Resource Management* (National Academy Press, Washington, DC, 1982).
11. N. C. Matalas, in *Global Change and Water Resources Management* (Water Resources Update No. 112, Universities Council on Water Resources, Carbondale, IL, 1998).
12. K. E. Schilling, E. Z. Stakhiv, in *Global Change and Water Resources Management* (Water Resources Update No. 112, Universities Council on Water Resources, Carbondale, IL, 1998).
13. J. R. Stedinger, D. Pei, T. A. Cohn, *Water Resour. Res.* **21**, 665 (1985).
14. Z. W. Kundzewicz, L. Somlyódy, *Water Resour. Manage.* **11**, 407 (1997).
15. P. C. D. Milly, K. A. Dunne, A. V. Vecchia, *Nature* **438**, 347 (2005).
16. J. C. Knox, *Quatern. Sci. Rev.* **19**, 439 (2000).
17. P. C. D. Milly, R. T. Wetherald, K. A. Dunne, T. L. Delworth, *Nature* **415**, 514 (2002).
18. Z. W. Kundzewicz et al., *Hydrol. Sci. J.* **50**, 797 (2005).
19. R. Seager et al., *Science* **316**, 1181 (2007).
20. IPCC, in *Climate Change 2007: Mitigation of Climate Change, Contribution of WG3 to AR4*, B. Metz et al., Eds. (Cambridge Univ. Press, New York, 2007), pp. 1–24.
21. A. Maass et al., *Design of Water-Resource Systems: New Techniques for Relating Economic Objectives, Engineering Analysis, and Government Planning* (Harvard Univ. Press, Cambridge, MA, 1962).
22. M. Reuss, *J. Water Resour. Plann. Manage.* **129**, 357 (2003).
23. E. L. Miles et al., *Proc. Natl. Acad. Sci. U.S.A.* **103**, 19616 (2006).
24. M. E. Moss, *Water Resour. Res.* **15**, 1797 (1979).
25. E. Ehrlich, B. Landy, *Public Works, Public Wealth* (Center for Strategic and International Studies Press, Washington, DC, 2005).
26. United Nations General Assembly, *U.N. Millennium Declaration*, Resolution 55/2 (2000).

MATERIALS SCIENCE

Sweet, Hairy, Soft, and Slippery

Synthetic polymer lubricants inspired by biological systems may be the key to water-based lubrication.

Seunghwan Lee and Nicholas D. Spencer

Water, together with surfaces containing sugar chains, forms the basis of all biological lubrication systems, from the slithering of a snail to the passage of food along the digestive tract. Yet humans have typically lubricated their machines with oils and fats. Understanding of biological lubrication has now advanced to the point where these principles can be applied to systems of technological importance using synthetic polymers.

Water lubrication is used in some niche applications such as reservoir pumps, and oil-water emulsions are often used for metal cutting, taking advantage of the effectiveness of water as a coolant. In the mining industry, hydraulic fluids are frequently based on water so as to exclude flammable materials from underground working areas. Interest in water lubrication is also high in the food, textile, and pharmaceutical industries, where product contamination by oil is a concern (1).

Water on its own is, however, generally a poor lubricant, and unlike oil, its viscosity does not rise substantially with pressure. This property is essential to the mechanism by which oils can form a lubricating film in high-pressure, nonconformal contacts of hard materials such as gears or ball bearings (1).

The low viscosity of water at high pressures can be overcome by biological lubricant additives, usually glycoproteins, in which large numbers of sugar chains are bound along a protein backbone. For example, mucins are found in most parts of the human body that need lubricating, such as eyes and knees (2). These molecules probably aid lubrication both via their intrinsic viscoelastic properties in solution (3) and via their behavior when adsorbed on the sliding surfaces. The characteristic bottlebrush structure of the molecules is crucial to this mechanism: The hydrophilic sugars immobilize large amounts of water within the contact region, while the backbone interconnects to other bottlebrushes or to a surface. Hierarchically structured, sugar-based bottlebrushes also play a key role in the mechanical properties of cartilage (see the figure) (4).

One biomimetic approach is to decorate the sliding surfaces with a high density of brush-forming polymer chains. Klein *et al.* have shown that when two mica surfaces bearing polymer brushes are rubbed past each other under compression in "good solvents," the interfacial friction forces lie below the detection limit (5, 6). The remarkable lubricating effect of such hairy polymer layers is ascribed to interchain repulsion, which leads to the incorporation of large quantities of solvent. The resulting fluid-like cushioning layer on the surface can sustain the externally applied pressure, thereby lowering the friction forces (7). This behavior, first observed for polystyrene chains in toluene, is a result of the interplay between the polymer and solvent

Learning from biology. Proteoglycan aggregate is a natural hierarchical bottlebrush that plays an important structural role in cartilage. Analogous synthetic systems could be used for water-based lubrication of technical applications such as sliding contacts or bearings.

rather than an intrinsic property of either component. Thus, it can also be observed for hydrophilic polymers such as polyethylene glycol (PEG) in water (8–10).

How high a pressure can brush-like polymer layers withstand? With increasing pressure, the compression and interpenetration of opposing polymer brushes also increase, even in good solvents; the disruption of the brushes, and thus the onset of substantial friction forces, occurs at pressures between 0.01 and 1 MPa, corresponding to pressures in contacts between internal organs or non–load-bearing mammalian joints (11). Nonetheless, Müller *et al.* attached dense PEG brushes to steel and glass surfaces sliding in water to investigate their lubricating properties under conditions approaching those of industrial bearings (12). They found

that the formation of a lubricating film was substantially enhanced by the brushes at contact pressures as high as 0.5 GPa (conditions where water alone cannot form a lubricating film), but the brush layer became detached during sliding contact, and direct contacts between peaks in the surface roughness could not be suppressed completely.

Another important characteristic of natural tribological systems is that they usually involve soft surfaces, as exemplified by slugs, eyes, tongues, and cartilage-coated articular joints. In response to external loads, such soft surfaces deform elastically and increase the contact area, resulting in a relatively low contact pressure. This is why liquids whose viscosity increases only slightly with pressure, such as water, can form lubricating films in soft contacts. This property sparked an extensive study of the aqueous lubrication of elastic polymers (rubbers or elastomers) (13) and has led to applications, for example, in tires, seals, windshield wipers, and biomedical implants.

The synergistic combination of hairy polymers and soft surfaces for water lubrication has, however, been investigated only recently. For example, end-grafted PEG chains can greatly enhance the water lubrication of silicone-rubber surfaces, especially at low speeds (14). In contrast, short-chain surfactants only led to minor reductions in friction in comparison to brush-like systems (15). Gong *et al.* have gone a step further in biomimicry by using hydrogels for aqueous lubrication. In addition to being soft, hydrogels allow water

The authors are in the Laboratory for Surface Science and Technology, Department of Materials, ETH Zürich, 8093 Zürich, Switzerland. E-mail: spencer@mat.ethz.ch

to permeate into the network, as does cartilage. Among the many hydrogels that have been investigated, those with brush-like polymer chains at their surfaces are the most effective for aqueous lubrication (*16*).

The combination of brushes with soft surfaces is clearly a key aspect of biological lubrication. However, it is less clear whether carbohydrates possess any specific or unique properties that are absent in other, synthetic brush-forming hydrophilic polymer chains. Moreover, natural lubricant additives appear to form hierarchical bottlebrush structures, such as that shown in the figure, more readily than the synthetic water-soluble brushes that have been investigated to date (*9*, *12*, *14*). The role of both the composition and struc-

ture of sugar-based, bottlebrush-structured molecules in natural lubrication thus needs to be clarified.

Although the feasibility of aqueous lubrication with elastomers and hydrogels has been established, the poor mechanical properties and wear resistance of soft materials have been limiting factors for applications. However, it has been shown that the mechanical properties of hydrogels can be improved to a similar level as those of elastomers (*17*). Thus, there is hope that hydrogels can be used in applications where mechanical strength is required. The practical implementation of biomimetic, aqueous lubrication approaches may become a reality in the not-too-distant future.

References
1. D. Dowson, *History of Tribology* (Professional Engineering Publishing, London, 1998).
2. R. Bansil *et al.*, *Annu. Rev. Physiol.* **57**, 635 (1995).
3. J. Celli *et al.*, *Biomacromolecules* **6**, 1329 (2005).
4. L. Han *et al.*, *Biophys. J.* **92**, 1384 (2007).
5. J. Klein *et al.*, *Nature* **370**, 634 (1994).
6. J. Klein, *Proc. Inst. Mech. Eng. Part J* **220**, 691 (2006).
7. P. G. De Gennes, *Macromolecules* **13**, 1069 (1980).
8. U. Raviv *et al.*, *Langmuir* **18**, 7482 (2002).
9. T. Drobek, N. D. Spencer, *Langmuir* 10.1021/la702289n (2007).
10. M. Kobayashi *et al.*, *Soft Matter* **3**, 740 (2007).
11. J. Klein *et al.*, *Acta Polym.* **49**, 617 (1998).
12. M. Müller *et al.*, *Tribol. Lett.* **15**, 395 (2003).
13. D. F. Moore, *The Friction and Lubrication of Elastomers* (Pergamon, Oxford, 1972).
14. S. Lee, N. D. Spencer, *Tribol. Int.* **38**, 922 (2005).
15. S. Lee, N. D. Spencer, *Lubric. Sci.* **20**, 21 (2008).
16. J. P. Gong, *Soft Matter* **2**, 544 (2006).
17. Y. Tanaka *et al.*, *Prog. Polym. Sci.* **30**, 1 (2005).

IMMUNOLOGY

The Toll of Cathepsin K Deficiency

An enzyme that is important to bone metabolism also acts in immune cell responses to pathogen DNA.

Arthur M. Krieg and Grayson B. Lipford

Cathepsins are lysosomal proteases that cleave their substrates with relatively low specificity. Cathepsins L and S process antigens for presentation to T cells, and are thus critical for developing adaptive immunity and long-lasting protection from specific pathogens (*1*). However, no cathepsin had been known to participate in innate immunity, a more general and nonspecific response to infection. On page 624 in this issue, Asagari *et al.* (*2*) reveal that cathepsin K, known for its importance in bone resorption, is important in the innate immune response to pathogen DNA. This presents an interesting example of molecules and mechanisms that are shared between the skeletal and immune systems, and raises the issue of whether treatment of conditions in one system could affect functions of the other.

Toll-like receptors (10 are expressed in humans) have been highly conserved throughout more than 400 million years of evolution. They are expressed in immune cells and enable the innate immune system to detect pathogens and, to an extent, distinguish whether they are extracellular or intracellular. Extracellular pathogens are detected by cell surface Toll-like receptors that recognize pathogen-expressed molecules such as diacyl and triacyl lipopeptides and lipopolysaccharides. In contrast, nucleic acids of intracellular

pathogens are detected by Toll-like receptors present in endosomal-lysosomal cellular compartments. Of these intracellular Toll-like receptors, the best understood is TLR9, which is specifically activated by unmethylated CpG motifs prevalent in bacterial and viral DNA.
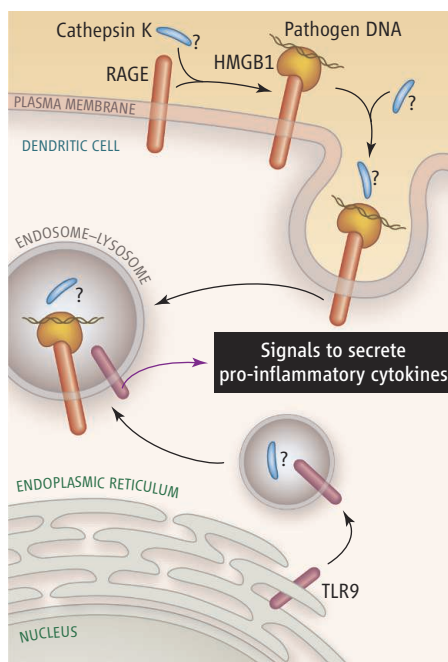
Possible points of action. Cathepsin K could affect the innate immune response to pathogen DNA by compromising TLR9 signaling at various points. TLR9 associates with the RAGE-HMGB1-DNA cofactor complex and elicits pro-inflammatory signals in response to infection.

Unlike other intracellular Toll-like receptors, TLR9 uses a cofactor called high-mobility group box 1 (HMGB1). This protein binds to both microbial DNA and the receptor for advanced glycation end-products (RAGE). This DNA-protein complex ultimately stimulates TLR9 activation (*3*).

Asagari *et al.* observed that NC-2300, a relatively specific small-molecule inhibitor of cathepsin K, suppressed bone resorption by osteoclasts in vitro and in vivo, and reduced bone demineralization in a rat arthritis model. Surprisingly, NC-2300 also had an anti-inflammatory effect, reducing paw swelling. In another autoimmune (encephalomyelitis) disease model, cathepsin K–deficient mice developed milder disease than wild-type mice, suggesting a more general function of the enzyme in regulating inflammation (NC-2300 was not tested in this model). NC-2300 did not inhibit antigen processing in bone marrow–derived macrophages, indicating that cathepsin K does not share this function with cathepsins L and S. However, NC-2300 nearly eliminated the TLR9-induced secretion of pro-inflammatory cytokines by bone marrow–derived dendritic cells, without affecting the immune responses of other Toll-like receptors to natural or synthetic ligands.

There are several possible explanations for decreased TLR9 activity in cathepsin K–compromised cells (see the figure). TLR9, a cofactor, or endosomal-lysosomal enzymes could be a direct substrate of cathepsin K. The enzyme could also act indirectly by activating

The authors are with the Coley Pharmaceutical Group, 93 Worcester Street, Wellesley, MA 02481, USA. E-mail: akrieg@coleypharma.com

or deactivating specific endosomal-lysosomal enzymes that control interaction between TLR9 and CpG DNA.

Cathepskin K could also be important for the transit of TLR9 to the endosomal-lysosomal compartment. TLR9 signaling is sensitive to its position within the endosomal-lysosomal network (bacterial CpG-A and CpG-B oligodeoxynucleotides signal from early versus late compartments, respectively, affecting dendritic cell activation). Toll-like receptors move from the endoplasmic reticulum to the endosomal-lysosomal network by a yet undiscovered trafficking signal and/or chaperone carrier. Chaperoning could be analogous to that needed for major histocompatibility complex class II molecules (which present antigens at the cell surface to T cells) to pass from the endoplasmic reticulum to the lysosome

before antigen loading and passage to the cell surface. During this transport, cathepsins L, F, S, and V are important for the chaperoning role of a protein called the invariant chain. By analogy, cathepsin S has been proposed to control the function of CD1d, a glycoprotein that presents lipid antigens to T cells. Cathepsin S may influence the transport of CD1d by cleaving key proteins in the endosomal-lysosomal network.

One important question is the degree to which the immune effect of cathepsin K is limited to TLR9. Is the apparent role of cathepsin K in regulating the autoimmune diseases examined by Asagiri *et al.* consistent with a simple TLR9 inhibition mechanism? TLR9 agonists may contribute to both autoimmune disease models examined in the study and TLR9-deficient mice are partially resistant to

experimental autoimmune encephalomyelitis (*4*, *5*). Nevertheless, it remains to be determined whether there is any pathogenic role for cathepsin K in human autoimmune or inflammatory diseases. The protease activity of cathepsin K is thought to be relatively nonspecific, including strong collagenase, elastase, and gelatinase activities, and how this translates into such an apparently specific effect on TLR9 activation remains to be elucidated.

References
1. L. C. Hsing, A. Y. Rudensky, *Immunol. Rev.* **207**, 229 (2005).
2. M. Asagiri *et al.*, *Science* **319**, 624 (2008).
3. J. Tian *et al.*, *Nature Immunol.* **8**, 487 (2007).
4. A. M. Krieg, J. Vollmer, *Immunol. Rev.* **220**, 251 (2007).
5. M. Prinz, *J. Clin. Invest.* **116**, 456 (2006).

10.1126/science.1154207

MATERIALS SCIENCE

Glass Surfaces Not So Glassy

Nanoscale indentations on a polymer glass relax rapidly, indicating the presence of a highly mobile liquid surface layer.

J. R. Dutcher and M. D. Ediger

In many materials, atoms or molecules on surfaces can be more mobile than particles in the interior. For example, the diffusion of an atom along a single crystal surface is typically much faster than diffusion inside the crystal. The control of layer-by-layer growth of crystalline materials for semiconductor lasers and other devices is possible because of a sophisticated understanding of this process. Many technologies also rely on noncrystalline materials such as glass (a solid without the regular packing of a crystal). Is this same type of mobility present on the surface of a glass? On page 600, Fakhraai and Forrest (*1*) report elegant experiments establishing that a polymer glass surface can be many orders of magnitude more mobile than the interior. But their results cannot be explained by atoms or molecules skittering across a surface. Rather, they find that a liquid-like layer, at least several nanometers thick, exists at the surface of a polymer glass.

Fakhraai and Forrest measure surface mobility by preparing polystyrene films with well-defined nanoindentations on the surface
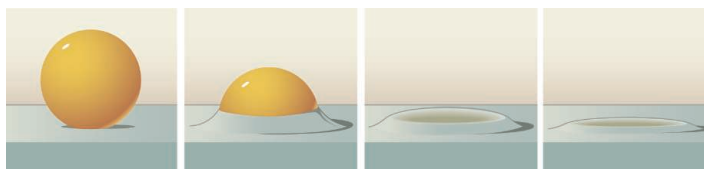
Watching glass relax. Gold particles are deposited, allowed to partially embed in the glass surface, and then gently removed. Fakhraai and Forrest used atomic force microscopy to image the filling of the nanoindentations over time at various annealing temperatures.

(see the figure). Gold spheres (with a nearly uniform 20-nm diameter) are deposited on a flat polystyrene surface at room temperature, well below the glass transition temperature T_g where polystyrene becomes a solid. Annealing the polymer films above T_g allows the gold spheres to sink a few nanometers into the surface. At room temperature, mercury dissolves away the gold spheres, leaving behind hemispherical nanoindentations about 5 nm deep. Surfaces prepared in this manner are then annealed below T_g for various periods of time, and atomic force microscopy is used to image the filling of the holes.

The most striking observation by Fakhraai and Forrest is that filling of the nanoindentations on the polystyrene glass surface is essentially complete on relatively short time scales. At 20 K below T_g the process takes a few minutes, whereas at 100 K below T_g the holes fill in a few weeks. These times are much faster (by a factor of ~10^9 at $T = T_g - 100$ K) than

comparable relaxation processes in bulk polystyrene. Additionally, the surface relaxation process has a much weaker temperature dependence than bulk relaxation, and the two processes appear to merge slightly above T_g.

Owing to the wide application of polymer glasses, researchers have been interested in understanding the dynamics of these surfaces (*2*, *3*). Prior studies measured the relaxation of a free surface upon which a pattern has been templated (*4*, *5*). However, the present work uniquely combines a reproducible deformation with a simple geometry, a penetration depth of only a few nanometers (necessary for probing only the surface), and a gentle method for removing the template from the sample. Although previous studies have provided evidence for enhanced mobility at the surface of polymer glasses (*6*), this is the first quantitative characterization of the surface relaxation time over a wide temperature range. The merging of the surface and bulk relaxations near T_g explains why previous surface-sensitive experiments performed close to T_g show nearly bulk dynamics.

The work of Fakhraai and Forrest highlights a critical difference between mobility at crystal surfaces and at glass surfaces. Because the polymer segments in these experiments are connected to each other in long chains that extend

J. R. Dutcher is in the Department of Physics, University of Guelph, Guelph, Ontario N1G 2W1, Canada. M. D. Ediger is in the Department of Chemistry, University of Wisconsin, Madison, WI 53706, USA. E-mail: ediger@chem.wisc.edu

20 nm or more into the glass, the nanoindentations cannot be filled by polymer segments skittering across the surface; each segment at the surface is chemically tethered to other segments that are below the surface. Instead, mobility at the surface of a glass can be viewed as the motion of a thin liquid-like layer that responds to surface tension, filling the nanoindentations to minimize surface area. These experiments and others indicate that the range of enhanced surface mobility is at least several nanometers (and perhaps more than 10 nm) (6); for comparison, the polymer segments are roughly 1 nm. A liquid-like layer of several nanometers is reasonable given that molecular motions near T_g are highly cooperative. That is, the packing is so tight that one polymer segment can move only when a large number of neighbors collectively adjust their positions (7). The size of this group of segments has been estimated to be several nanometers (8), and it makes sense that the presence of a surface would perturb motions over a length scale at least this large (9). Whether this viewpoint can yield a predictive theory of glass surface mobility is unknown.

The surface mobility of polymer glasses is relevant for understanding adhesion, friction, and instabilities in thin polymer films and nanostructures used in advanced lithography (10). If strategies can be devised to decrease surface mobility, photoresists could be used to transfer patterns on even smaller length scales. Surface mobility is also likely to be important in a wide range of nonpolymeric glasses (11, 12). Given the enormous magnitude of this effect in polystyrene, the surface mobility of glasses likely has important consequences that have not yet been considered.

References
1. Z. Fakhraai, J. A. Forrest, *Science* **319**, 600 (2008).
2. J. L. Keddie *et al.*, *Europhys. Lett.* **27**, 59 (1994).
3. C. B. Roth, J. R. Dutcher, in *Soft Materials: Structure and Dynamics,* J. R. Dutcher, A. G. Marangoni, Eds. (CRC, New York, 2004), pp. 1–38.
4. T. Kerle *et al.*, *Macromolecules* **34**, 3484 (2001).
5. E. Buck, K. Peterson, M. Hund, G. Krausch, D. Johannsmann, *Macromolecules* **37**, 847 (2004).
6. C. J. Ellison, J. M. Torkelson, *Nat. Mater.* **2**, 695 (2003).
7. E. R. Weeks *et al.*, *Science* **287**, 627 (2000).
8. U. Tracht *et al.*, *J. Magn. Reson.* **140**, 460 (1999).
9. P. Scheidler, W. Kob, K. Binder, *J. Phys. Chem. B* **108**, 6673 (2004).
10. K. Yoshimoto *et al.*, *J. Chem. Phys.* **122**, 144712 (2005).
11. S. F. Swallen *et al.*, *Science* **315**, 353 (2007); published online 6 December 2006 (10.1126/science.1135795).
12. R. C. Bell *et al.*, *J. Am. Chem. Soc.* **125**, 5176 (2003).

10.1126/science.1155120

MOLECULAR BIOLOGY

The Art of Assembly

Francis Szoka

Advances in delivering small interfering RNAs to specific tissues may bring these nucleotides closer to reality as therapeutic agents.

Small RNAs were chosen as the "Breakthrough of the Year" molecule for 2002 (1). One of these, small interfering RNA (siRNA), which is 20 to 23 nucleotides in length, can base pair with a target messenger RNA (mRNA) sequence and direct its degradation, thus blocking production of the encoded protein. Although siRNAs have surpassed expectations when used in experiments to alter gene expression, the challenge of turning them into effective drugs has been the lack of an efficient in vivo delivery system. On page 627 in this issue (2), Peer *et al.* describe a strategy that offers cautious optimism for siRNA becoming a therapeutic reality to treat human disease in the coming decade.

Although small by nucleic acid standards, siRNAs are large compared to most drugs. A 22-nucleotide siRNA has a molecular weight of about 15,000, about 50 times that of a typical drug. Because they also have a strong negative charge, siRNAs cannot readily cross biological membranes and enter cells. Furthermore, siRNAs are metabolized in the blood, requiring a chemical modification or a specific formulation solution to prevent their degradation during delivery.

Despite these limitations, delivering siRNA by attaching it to a lipid modestly reduced gene expression in mouse hepatocytes (3). Substantial improvement in siRNA delivery in nonhuman primates was achieved by packaging the siRNA in a cationic liposome that was coated with polyethylene glycol polymer chains (4). At a single dose of 2.5 mg of siRNA per kilogram of body weight, this approach reduced expression in the liver of the cholesterol-carrying protein apolipoprotein B by more than 90%, and also decreased serum cholesterol (4). Both effects persisted for more than 2 weeks. By comparison, a 2-week supply of an oral cholesterol-reducing drug is about 2 mg per kilogram of body weight.

Most in vivo siRNA delivery studies have been directed to tumor tissue or to the liver, so the study by Peer *et al.* targeting leukocytes is an exciting new development. Promising delivery approaches described by this group and others have a unifying theme: attaching a targeting ligand—either a carbohydrate (5), small molecule (6), antibody (2), or the mole-

The author is in the Department of Biopharmaceutical Sciences and Pharmaceutical Chemistry, University of California, San Francisco, CA 94143–0446, USA. E-mail: szoka@cgl.ucsf.edu

Performance of siRNA carriers with proven intravenous delivery capabilities					
Species tested	**Carriers of siRNA**	**Ligand attached to carrier**	**Target of siRNA**	**Effects**	**Adverse side effects**
Primate (4)	Liposome–polyethylene glycol complex	None	Apolipoprotein B (liver)	Decreases serum cholesterol	Transient increase in serum transglutaminases
Mouse (5)	Polymer	Carbohydrate	Apolipoprotein B (liver)	Decreases serum cholesterol	Transient increase in serum transglutaminases
Primate (7)	Polymer	Transferrin	Ribonucleotide reductase (tumor)	Inhibits M2 subunit of ribonucleotide reductase	Immune response to transferrin; increase in inflammatory cytokines
Mouse (6)	Liposome–protamine complex	Anisamide	Epidermal growth factor receptor (tumor)	Blocks epidermal growth factor activity and tumor proliferation	Transient increase in serum transglutaminases
Mouse (2)	Liposome–hyaluronan–protamine complex	Antibody to β_7 integrin	Cyclin D1 (leukocytes)	Reduces leukocyte proliferation and inflammation	None reported

cule transferrin (7)—to a delivery complex consisting of a siRNA that is compacted by either protamine or a cationic polymer, and stabilized by polyethylene glycol or hyaluronan (see the table). The resulting nanoparticles are under 150 nm in diameter, a size that enables a sufficiently long circulation time so that the complex can reach the target site.

Leukocytes and other hematopoietic cells are finicky when it comes to internalizing cell surface–bound antibodies, and consequently, delivering liposomes decorated with antibodies to immune cells has been a challenge (8). Peer et al. covalently attached two molecules to 100-nm noncationic liposomes in sequence—first, hyaluronan, and then an antibody to β_7 integrin, an adhesion molecule expressed by leukocytes that traffic to gut tissue. To form the siRNA delivery complex, they then combined the antibody-hyaluronan-liposome with a siRNA (directed against mRNA encoding cyclin D1, a regulator of the cell division cycle) that was compacted with protamine. The final purified liposome complex was injected intravenously into mice.

The approach worked. The siRNA silenced cyclin D1 expression in leukocytes, suppressed leukocyte proliferation, and reversed intestinal inflammation normally triggered by excessive leukocyte activity in a mouse model of colitis.

The work by Peer et al. is a technical tour de force. The sequential assembly process involves two chemical couplings that, if not done properly, can potentially generate N-acylurea groups on the hyaluronan. When N-acylurea groups are hydrolyzed into the ethylurea compound, they have an anti-inflammatory effect (9). Extensive experimental controls by Peer et al. provide a high level of assurance that the anti-inflammatory effect observed results from decreased cyclin D1 expression by siRNA that is delivered to leukocytes.

This recent emergence of multiple targeted carrier systems bodes well for the future of in vivo siRNA delivery. However, there is room for improvement. The current protocols require many components and multiple assembly steps to package siRNA in a delivery system. In addition, if the siRNA carrier formulation is administered on multiple occasions, it may generate antibodies against the targeting ligand attached to the carrier and perhaps to some of the packaging elements such as protamine. For instance, antibodies against human transferrin were elicited when a polymer formulation using transferrin as the targeting ligand was reinjected into nonhuman primates (7). The capacity of siRNA delivery systems to induce undesired immune reactivity (10) requires careful examination of immune responses to the systems as targeting approaches mature. Nonetheless, the studies by Peer et al. and others highlight the types of targeted systems that may be optimized to provide a robust siRNA delivery to the liver, tumors, and now to the hematopoietic system.

References
1. J. Couzin, Science 298, 2296 (2002).
2. D. Peer, E. J. Park, Y. Morishita, C. V. Varman, M. Shimaoka, Science 319, 627 (2008).
3. J. Soutschek et al., Nature 432, 173 (2004).
4. T. S. Zimmermann et al., Nature 441, 111 (2006).
5. D. B. Rozema et al., Proc. Natl. Acad. Sci. U.S.A. 104, 12982 (2007).
6. S. D. Li, Y. C. Chen, M. J. Hackett, L. Huang, Mol. Ther. 16, 163 (2008).
7. J. D. Heidel et al., Proc. Natl. Acad. Sci. U.S.A. 104, 5715 (2007).
8. C. O. Noble et al., Expert Opin. Ther. Targets 8, 335 (2004).
9. B. Ruiz-Perez et al., J. Infect. Dis. 188, 378 (2003).
10. J. T. Marques, B. R. Williams, Nature Biotechnol. 23, 1399 (2005).

10.1126/science.1154253

MATERIALS SCIENCE

Nanowires in Nanoelectronics

Connecting circuit layers with nanowires and nanotransistors may bring about a paradigm shift in microchip design.

David K. Ferry

For almost four decades, progress in microchips has followed Moore's (1) famous dictum that the transistor density would double roughly every 18 months. This steady progress has brought us to the point today where leading-edge chips have transistors whose critical dimension is only about 100 atoms long. Clearly, this evolution cannot continue down this same path much longer. Recognition of the impending "end of the road" has led many to seek an alternative to the ubiquitous silicon transistor, hoping thereby to revolutionize the industry that has fueled the massive information revolution since World War II (2, 3). Among the promising candidates are nanowires and transistors made from them (4–7). These nanowires have been grown as carbon nanotubes or from silicon, as well as a variety of other semiconductors. As with most new technologies, there are

The author is in the Department of Electrical Engineering, Arizona State University, Tempe, AZ 85287, USA. E-mail: ferry@asu.edu

Going vertical. A schematic, conceptual view of the introduction of vertical nanowires on a microchip. The bottom layer is a chip layout drawing; the various colors represent interconnection levels on the chip. The nanowires can reach from the chip level to higher-lying interconnects or they can reach between various metal layers.

remarkable expectations for the usefulness of these nanowires. In reality, they are not likely to replace the silicon transistor, but they may well provide the paradigm shift that will extend Moore's "law."

To understand how this paradigm shift must occur, we need to understand the driving force for Moore's Law. It is a trend that does not derive from physical science but from economics. Transistors are laid out on the microchip in a planar fashion, much like houses in a modern southwestern city. According to

Intel, the latest 45-nm microprocessor (with a gate length—the critical dimension in the direction of current flow—of ~22 nm) has about 410 million transistors in 107 mm^2, or each transistor occupies a square of silicon real estate that is roughly 500 nm on a side.

Originally, Moore's Law was driven by three factors: (i) reducing the transistor size (and therefore the square of silicon upon which it sits), (ii) increasing the size of the microchip itself, and (iii) circuit cleverness (by which the number of transistors needed to perform a function could be reduced with consequent savings in silicon real estate). As the number of transistors increased, the number of functional units in each chip could be

CREDIT: MICROCIRCUIT LAYOUT DIAGRAM ADAPTED FROM L. T. CLARK/ARIZONA STATE UNIVERSITY

increased, thus increasing the computing power per chip. Because the basic cost of manufacturing the microchip has not dramatically increased over these four decades, the cost per functional unit, and the cost of computing power, has gone down exponentially. It is this economic argument, the cost of silicon real estate, that drives Moore's Law.

For the past decade, however, the size of the microchip has remained roughly constant, so that factors (i) and (iii) have become more important. Today, as pointed out above, we are approaching the atomic limit on critical size. We are left, therefore, with the conclusion that it is factor (iii) that will have to provide the continuity to Moore's Law. What does this say about the role of nanowires?

Although transistors have very short gate lengths (the critical dimensions mentioned above), they have much larger widths in order to provide the current necessary to make the circuits work. If we are to replace the current planar transistor with nanowires, we will have to use a great many such devices in parallel to provide this current. But, we must satisfy the above economic driving forces, and effectively use the overall silicon real estate. This leads to a geometric argument that says that nanowires in the plane will not effectively compete with novel transistors such as the "fin" field effect transistor, or "finFET," Basically, the finFET is a vertically oriented Si "fin" in which transistor structures can be placed on both sides (8, 9), and even on the top (10), of this fin. A properly configured finFET more effectively uses silicon real estate. Consequently, there does not seem to be much of a role for nanowires laid horizontally on the Si surface (11, 12). However, nanowires can be grown vertically, and this growth can be initiated on a wide variety of substrates (4).

As mentioned, currently the transistors are placed in a planar array, but they are overlaid with a great many metal lines, with nine or more levels of metal (worse than any freeway interchange). These metal lines provide power, clock signals to synchronize the switching of the transistors, and various interconnections between the different functional blocks of the chip. Connections from these metal lines to the transistors are made by downwardly reaching metal fingers called "vias." When we can no longer reduce the transistor size, we enhance the use of Si real estate by moving vertically.

Our third Moore's Law factor—cleverness—can be increased by replacing some of these vias with vertical nanowire transistors (13) (see the figure). These vertical transistors can reach from the silicon to a metal wire or even between different levels of metal wire. Moreover, we can begin to think about creating reconfigurable architectures in which the connections between different functional blocks are changed by switching just a few of these vertical transistors. Thus, we begin to create real three-dimensional architectures in a different manner from the traditional approach of stacking chips (14).

If we are to use these vertical transistors for more effective architectures, then we have to change how we go about microchip design. Today, this chip design is done with automated transistor layout programs that optimize the planar design placing of the various functional blocks and minimize the necessary interconnections (in the metal layers). To change to reconfigurable architectures, we need switchable interconnections based on vertical transistors, and device physicists will have to work with circuit designers to achieve this. These new opportunities for nanowires to extend Moore's Law may well force this paradigm shift.

References

1. G. Moore, *Electronics* **5**, 114 (1965).
2. C. Evans, *The Micro Millenium* (Washington Square Press, New York, 1979).
3. T. Forester, *The Microelectronics Revolution* (MIT Press, Cambridge, MA, 1982).
4. D. R. Bowler, *J. Phys. Cond. Matt.* **16**, R721 (2004).
5. W. Cui, C. M. Lieber, *Science* **291**, 851 (2001).
6. J. Appenzeller *et al.*, *Phys. Rev. Lett.* **92**, 226802 (2004).
7. T. Bryllert *et al.*, *IEEE Electron. Dev. Lett.* **27**, 323 (2006).
8. D. J. Frank *et al.*, *IEDM Tech. Dig.* **1992**, 553 (1992).
9. D. Hisamoto *et al.*, *IEEE Trans. Electron Dev.* **47**, 2320 (2000).
10. H. S. Doyle *et al.*, *IEEE Electron Dev. Lett.* **24**, 263 (2003).
11. R. Chau *et al.*, *IEEE Nanotechnol.* **4**, 153 (2005).
12. D. K. Ferry, *Phys. Stat. Sol. (c)* **5**, 17 (2008).
13. G. S. Duesberg *et al.*, *Diamond Rel. Mat.* **13**, 354 (2004).
14. See, for example, M. Koyanagi *et al.*, *IEEE Trans. Electron Dev.* **53**, 2799 (2006).

CLIMATE

Food Security Under Climate Change

Food insecurity is likely to increase under climate change, unless early warning systems and development programs are used more effectively.

Molly E. Brown and Christopher C. Funk

Some of the most profound and direct impacts of climate change over the next few decades will be on agricultural and food systems. On page 607 of this issue, Lobell *et al.* (*1*) show that increasing temperatures and declining precipitation over semiarid regions are likely to reduce yields for corn, wheat, rice, and other primary crops in the next two decades. These changes could have a substantial impact on global food security.

M. E. Brown is at Science Systems and Applications, Inc., NASA Goddard Space Flight Center Biospheric Sciences Branch, Code 614.4, Greenbelt, MD 20771, USA. C. C. Funk is in the Climate Hazard Group, Geography Department, University of California, Santa Barbara, CA, 93106, USA. E-mail: molly.brown@nasa.gov; chris@geog.ucsb.edu

Since the 1990s, rising commodity prices and declining per capita cultivated area have led to decreases in food production, eroding food security in many communities (*2*). Many regions that lack food security rely on local agricultural production to meet their food needs. Primarily tropical and subtropical, these regions are substantially affected by both global climate variations and global commodity price fluctuations. Warming in the Indian Ocean (*3*) and an increasingly "El Nino–like" climate (*4*) could reduce main-season precipitation across parts of the Americas, Africa, and Asia (see the figure).

In food-insecure regions, many farmers both consume their product and sell it in local markets. This exposes farmers to climate variations, because when they produce less their income goes down while their costs go up to maintain basic consumption. Large-scale hunger can ensue, even when there is sufficient food in the market that has been imported from elsewhere.

National revenue can also be affected by large-scale droughts, which restrict the ability of countries with small budgets to purchase grain on the international market. Thus, recent large increases in grain prices reduce access to food for the poor, for example, in Tanzania, who compete for corn with ethanol producers and hog farmers in the United States. Finally, up to half of all malnutrition is driven by nonfood factors through diseases such as HIV/AIDS and malaria; the latter disease is
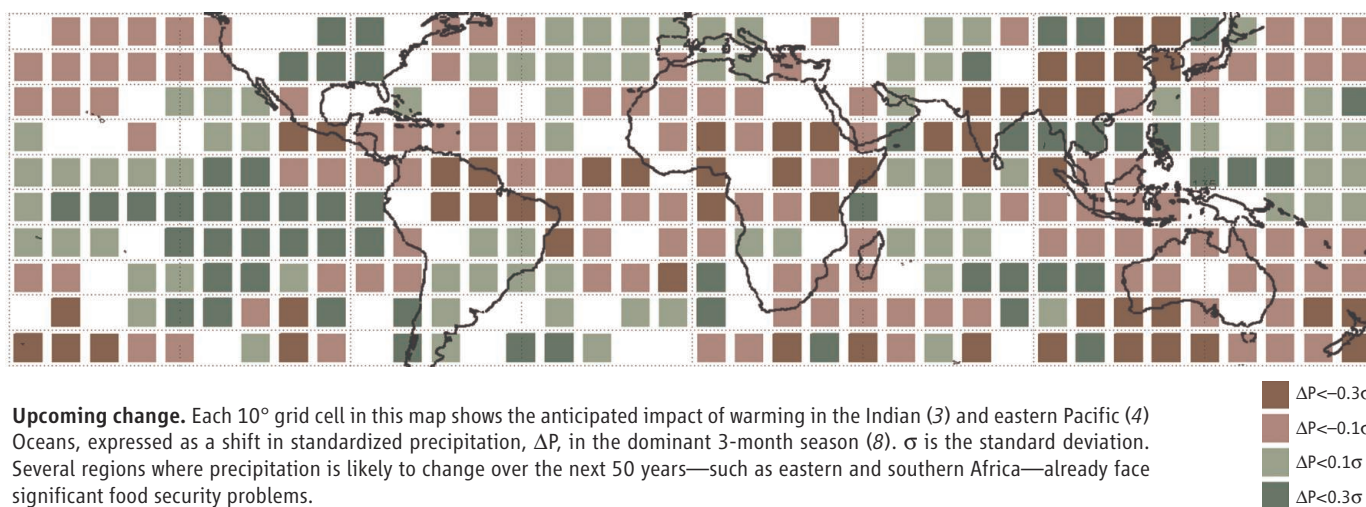
Upcoming change. Each 10° grid cell in this map shows the anticipated impact of warming in the Indian (*3*) and eastern Pacific (*4*) Oceans, expressed as a shift in standardized precipitation, ΔP, in the dominant 3-month season (*8*). σ is the standard deviation. Several regions where precipitation is likely to change over the next 50 years—such as eastern and southern Africa—already face significant food security problems.

ΔP<−0.3σ
ΔP<−0.1σ
ΔP<0.1σ
ΔP<0.3σ

likely to become more severe and widespread with warming temperatures.

Lobell *et al.* use crop models to calculate changes in agricultural production to 2030. The results show that climate change is likely to reduce agricultural production, thus reducing food availability. Identifying the impact of this reduced production will, however, be complicated by other changes. The latter include rising oil prices, the globalization of the grain market, and a structural change in demand for key food supplies due to increasing demand for biofuels and rising per-capita consumption in India and China. These changes have pushed up supply costs for staple foods by 40% or more in many food-insecure areas. Decoupling these effects to implement mitigation and adaptation programs will be difficult.

Climate change impacts on farmers will vary by region, depending on their use of technology. Technological sophistication determines a farm's productivity far more than its climatic and agricultural endowments. Food insecurity, therefore, is not solely a product of "climatic determinism" and can be addressed by improvements in economic, political, and agricultural policies at local and global scales. In currently food-insecure regions, farming is typically conducted manually, using a hoe and planting stick with few inputs. The difference between the productivity of these farms and those using petroleum-based fertilizer and pesticides, biotechnology-enhanced plant varieties, and mechanization is extreme (*5*). Not only will climate change have a differential effect on ecosystems in the tropics due to their already warmer climates, but also poor farmers in the tropics will be less able to cope with changes in climate because they have far fewer options in their agricultural system to begin with. These handicaps can be exacerbated by macro-economic policies that create disincentives for agricultural development,

such as agricultural subsidies in the United States and Europe and poorly implemented cash transfer programs (*6*).

The study by Lobell *et al*. suggests that communities can cope with climate change, for example, by switching from producing corn to producing sorghum, whose lower water requirements and higher temperature tolerances are better suited to a warmer and drier climate. However, this adaptation measure may be impossible to implement in many parts of the developing world. For example, it assumes markets for millet in regions where only maize is eaten, and technology and know-how about how to process and consume sorghum in maize zones. Communities may nevertheless be forced, as they are today, to consume what they produce regardless of cultural preferences.

Today, millions of hungry people subsist on what they produce. If climate change reduced production while populations increase, there is likely to be more hunger. However, it may still be possible to reduce world hunger through programs that feed the poor during crises and by investing in agricultural inputs such as fertilizer and improved varieties that can dramatically increase yields (*2*). Improved environmental monitoring and prediction systems can provide more effective early warnings, which may help governments to take action to preserve the thin agriculture production margins by which many make ends meet (*7*). Early warning systems involve extensive climate monitoring and prediction tools that could be used to enhance agricultural development programs. Crop insurance programs that are triggered by remote sensing data products may ensure farmer's livelihoods even in drought years. Investments in improved seeds and varietals and an augmented use of inorganic fertilizer (*2*, *6*) can increase yields. Improved local governance, reduced developed-world agricul-

tural subsidies, and more nuanced food aid policies that protect local markets could together produce rapid improvements in food access and availability, reducing hunger while providing for more people.

30% of farmers in developing countries are food-insecure; the work of Lobell *et al*. suggests that climate change may impact these undernourished communities by decreasing local yields while contributing to a global increase in commodity prices through significant global reduction in the production of corn, wheat, and rice. Despite these challenges, the very low agricultural productivity of food-insecure countries presents a great opportunity. Transform these agricultural systems through improved seed, fertilizer, land use, and governance, and food security may be attained by all.

References and Notes
1. D. B. Lobell *et al.*, *Science* **319**, 607 (2008).
2. Food and Agricultural Organization, *The State of Food and Agriculture*, 2007.
3. J. Verdin, C. C. Funk, G. Senay, R. Choularton, *Philos. Trans. R. Soc. B* **360**, 2155 (2005).
4. S. Solomon *et al.*, *Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the IPCC* (Cambridge Univ. Press, Cambridge, UK, 2007).
5. C. Thirtle, L. Lin, J. Piesse, *World Develop.* **31**, 1959 (2003).
6. OXFAM, *Causing Hunger: An Overview of the Food Crisis in Africa*, OXFAM briefing paper, July 2006; www.oxfam.org.uk/what_we_do/issues/conflict_disasters/bp91_hunger.htm.
7. M. E. Brown, C. C. Funk, G. Galu, R. Choularton, *EOS, Trans. Am. Geophys. Union* **88**, 381 (2007).
8. The figure shows statistically reformulated (*7*) CO$_2$ doubling scenario results from a multimodel ensemble (*9*) simulating changes between 2000 and 2050. Historical relationships in observed precipitation (*10*) are used to assess the probable impacts of the warming Indian (*3*) and eastern Pacific (*4*) Oceans.
9. G. A. Meehl *et al.*, *Bull. Am. Meteorol. Soc.*; 10.1175/BAMS-88-9-1383.
10. R. F. Adler *et al.*, *J. Hydrometeor.* **4**, 1147 (2003).

Insights into Phases of Liquid Water from Study of Its Unusual Glass-Forming Properties

C. Austen Angell

The vitrification of pure water is compared with that of molecular solutions rich in water, and gross differences are noted. Thermodynamic reasoning and direct observations on noncrystallizing nanoconfined water indicate that the glass transition in ambient-pressure water is qualitatively distinct from that found in the usual molecular liquids. It belongs instead to the order-disorder class of transition seen in molecular and ionic crystalline materials. The distinctive "folding funnel" energy landscape for this type of system explains the extreme weakness of the glass transition of water as well as the consequent confusion that has characterized its scientific history; it also explains the very small excess entropy at the glass transition temperature. The relation of confined water behavior to that of bulk is discussed, and the "fragile-to-strong" transition for supercooled water is interpreted by adding a "critical point–free" scenario to the two competing scenarios for understanding supercooled bulk water.

Many aspects of water's liquid-state behavior are unusual, and this includes the way in which it forms a glass. For most molecular liquids, particularly in small samples, rapid cooling generates a glassy solid rather than a crystalline phase. Vitrification in most cases has a clear thermodynamic signature—a rapid drop in heat capacity—as the translational and reorientational degrees of freedom, by which the liquid absorbs energy and flows, are frozen out. Water is exceptional in this regard in that its glass formation signature is very weak, so much so that the glass transition temperature T_g, typically assigned in the range from 120 to 160 K, has been controversial. In this review I argue that the various experimental and theoretical arguments about the glass transition in water can be reconciled in terms of an order-disorder transition that occurs in the "anomalous regime," 150 to 250 K (1–3). This order-disorder transition, which may include some weak first-order character, effectively steals the disordering heat capacity of the normal glass transition, leaving almost nothing to observe at the temperature predicted by T_g extrapolations from binary-solution glasses.

What Is the Glass Transition?

The glass transition is not a "transition" in the thermodynamic sense of the word (4). With the glass transition, everything depends on time, and so T_g can only be given a firm value when the time rate of change of temperature (heating or cooling) has been fixed. Then it is straightforward to say what the glass transition is, in terms of some basic property such as heat capacity. The glass transition is that phenomenon in which the heat capacity increases abruptly during heating,

Department of Chemistry and Biochemistry, Arizona State University, Tempe, AZ 85287, USA. E-mail: caa@asu.edu

as previously frozen degrees of freedom are excited. The heat capacity changes from a value characteristic of a solid (only vibrational degrees of freedom, often the Dulong and Pettit value) to the higher value characteristic of a liquid—and there is usually a big difference, making it easy to detect. For instance, in a glass-forming solution containing two parts of water and one part of its look-alike, hydrogen peroxide (H_2O_2, melting point $T_m = 271.5$ K), the heat capacity more than doubles at the glass transition (5). The same is found when the solution contains, in place of H_2O_2, the other water look-alike, hydrazine [N_2H_4, $T_m = 5.5°C$ (5)] (Fig. 1).

The glass transition also has flow consequences. Above its T_g, the substance, when stressed, flows like a liquid—but it flows very slowly, unless $T \gg T_g$. Below T_g (6–8), it bends elastically like a solid, and then splinters—unless time is first allowed to pass, when again it flows (8). Why? Well, the extra time that passed was equivalent to a lowering of the T_g. The glass transition occurs when (9) the condition

$$d\tau/dT \cdot dT/dt \ (= Q \cdot dT/dt) \approx 1 \qquad (1)$$

is met, where Q is the cooling rate, t is time, and τ is the "relaxation time".

In some liquids, those in the "fragile" class (10), the change in heat capacity at T_g happens very sharply and is complete in just a few degrees, because a fragile liquid is one whose relaxation time changes very rapidly, in "super-Arrhenius" fashion, with temperature. In non-fragile liquids, however, it may take tens of degrees, even hundreds in the case of "strong" liquids with high T_g values [this is also true for ordinary liquids studied on computer simulation time scales, e.g., 50°C for SPC-E water (11)]. The hydrazine solution in Fig. 1, for instance,

although having a very large heat capacity change, is not a fragile liquid (12). It requires some 20 K to complete its transition. It is a part of the controversy concerning water that some authors consider it to be a strong liquid near its T_g (13–17), whereas others have argued that it is very fragile (18, 19) and still others think it is both (14), depending on the temperature and the cooling rate.

The Problem with Water and Its T_g

The large change in heat capacity ΔC_p at T_g for $(H_2O)_2 \cdot H_2O_2$ (and the comparably large ΔC_p for $H_2O \cdot N_2H_4$) renders the T_g determinations in these cases quite unambiguous. However, in the phenomenon usually considered to be the glass transition of laboratory water (Fig. 1), the value of ΔC_p is tiny, only 2% of that for the above aqueous solutions (20, 21). When a glass transition is so feeble, it is quite difficult to detect. Experienced glass researchers have missed it completely (22, 23). Yue and the author (24) suggested such a weak effect might be rationalized as an "annealing pre-peak," the real T_g being inaccessible because of crystallization, but there are alternatives to consider. What other reason could there be for the difference between water and its solutions? We provide an answer below.

Note that the glassy water for which the above small value of ΔC_p was recorded was made from the liquid state by hyperquenching (25) and is called hyperquenched glassy (HQG) water. Other ways of making glassy water exist, some of them exotic (3). The most common method and the earliest reported (26)—and also the way glassy water forms in nature—is by vapor deposition on a cold surface. [Condensed on cosmic dust particles and agglomerated into comets, this form, called low-density amorphous (LDA) water, is believed to be the most abundant form of water in the universe.] Another method, more recent and quite exotic, is by compression of ordinary ice (I_h) until amorphized (27), followed by decompression with annealing. The latter process converts the initial high-density amorphization product HDA into LDA. The three forms, HQG and the two LDAs, are almost identical in structure, although the form obtained by the high-pressure route shows some differences in thermal behavior that are not removed by heating above 136 K, when structural differences between glasses of the same glass prepared by different routes would be expected to disappear (28). Each of these preparations manifests comparably feeble heat capacity jumps at the annealing-induced thermal event, usually close to 136 K (29, 30). But in the HDA → LDA case (29), a heat capacity scan, made at a rate slower by a factor of 120 than the others, found T_g at 124 K (without annealing), which is 12 K below the usual value and 17 K below the "annealing-free" (20 K/min) scan value, 141 K, reported in (20) and (24).

An important part of the problem in characterizing water through its glass transition is that

water is a very bad glass-former. This means that the glassy state does not survive heating much above the putative T_g of 136 K. Indeed, in Handa and Klug's low–scan rate study (29) it crystallized at 132 K. In a "standard" scan [20 K/min, see (6)], LDA crystallizes in the range 150 to 160 K, depending on the manner of preparation. Thus, there is a gap [or a "no-man's land" (31)] between 150 and 236 K (the limit of supercooling) in which no direct data are available for bulk, or even microscopic (~1 µm), water (although there are insights from nanoconfined water; see below).

In Fig. 1, what has been reviewed up to this point is summarized as the heat capacity changes at T_g for water and its solutions as well as for the higher-temperature liquid states of water and its solutions, both molecular and ionic. The behavior of water in the region immediately below its freezing point (5, 32, 33) is especially noteworthy. Although criticized as artifactual (34), it has recently been established as correct by bulk water (10-ml samples of extreme purity) studied to –30°C (35). The hydrazine solution data in Fig. 1 may be interpreted as how water would behave if it could not form an open tetrahedral network through hydrogen bonding.

The challenge brought into focus by Fig. 1 is to find out what happens in the "gap" for the case

Although the possibility of a weak first-order phase transition en route is not excluded (see below), we look for the behavior along a continuous path in C_p between the large and anomalously increasing value around 240 K, and the tiny value at 150 K.

Insights from Nanoscopic Water

The problem of how water should behave if crystallization did not occur was addressed some time ago by applying thermodynamic principles to data available from both below and above the gap (15). The results will be revisited below, but there is also a more direct route: by studying water in nanoscopic confinement. Water near surfaces often does not crystallize upon cooling, but only recently have the properties of such water been measured. The heat capacity of nanoscopically confined water was reported by the Oguni group (38, 39), who used measurements on water in nanoporous media of different pore sizes to separate the properties of "surface water" from those of "internal" water in their noncrystallizing smallest-pore samples. The "internal" component (38, 39) is shown in Fig. 2B, where the data are compared with the results of the thermodynamic analysis of Fig. 2A. Although some alternatives to Oguni's analysis may be considered, the heat

nomenon, not a glass transition. The only irreversibility observed in Oguni et al.'s data was a very feeble "glass transition" (not visible on the scale of Fig. 2B) that occurred at 160 K. Confinement seems to act like a pressure (40) or, as shown in (41), like a change of potential such as to displace a liquid-liquid critical point from negative to positive pressures, which then has the same effect. This implies that bulk water behavior at ambient pressure would be either sharper than Fig. 2B, like the second of the Fig. 2A curves, or weakly first-order, as discussed below.

In light of Fig. 2, the atypically small value of the C_p "jump" at T_g of HQG and other low-density forms of amorphous water can be readily understood. At 136 K, Fig. 2B shows that there is almost nothing of the liquid-phase excess heat capacity left to lose at a glass transition. Thus, although the glass transition arrests the structure, most of the change in heat capacity, and creation of disorder upon heating, occurs in the subpeak part of the 225 K transition. This transition converts water continuously from a "fragile" molecular liquid above the transition to a "strong" liquid below it, so that water is a "strong" liquid as it vitrifies, as claimed in (13, 14). This picture suggests behavior akin to classical network glasses, which lie at the "strong" extreme of glass-formers and are also characterized by small heat capacity signatures at T_g and Arrhenius temperature dependences of transport coefficients above it [see also (42)].

The form of the excess C_p of water seen in Fig. 2 is completely different from that of common glass-formers but resembles that of the classical order-disorder transition seen in superlattice alloys and rotator phases. In this respect, it is compatible with the much-discussed "second critical point" interpretation of water's anomalies (43–45) because a critical point, associated with the vanishing of the transition order parameter, is the most interesting outcome of these cooperative excitation (order-disorder) processes. However, it is not the necessary, nor the most common, outcome, as considered further below.

Order-disorder transitions were at the center of interest in condensed matter physics half a century ago, but they are not often discussed in the current literature. Nonetheless, water (at least at low pressures) can be

in which water does form the network but does not crystallize. The question is meaningful because we know that the glass can be produced from the ambient-pressure liquid by fast cooling (25, 36). Furthermore, liquid water can itself be generated (as tiny growing droplets from the supersaturated vapor) in this "gap" temperature range (37), and, in due course, it will be structurally characterized there as it is produced.

capacity revealed is remarkably similar to that given by the thermodynamic analysis at its more "smeared-out" limit, as shown in Fig. 2. When the heat capacity of ice is added to the thermodynamically deduced "excess heat capacity" of Fig. 2A, the agreement is almost quantitative. The striking feature of each curve is, of course, the peak at about 225 K. Note that this behavior is reversible; it is a liquid-state phe-

understood in terms of such a transition, as discussed below. However, it is useful to first consider a structural study by Mayer and co-workers (46) that suggests a rather abrupt crossover in the behavior of water during hyperquenching, which occurs in the vicinity of the heat capacity peak temperature of Fig. 2.

Mayer and co-workers used infrared spectroscopy to study the ion-pairing equilibrium in
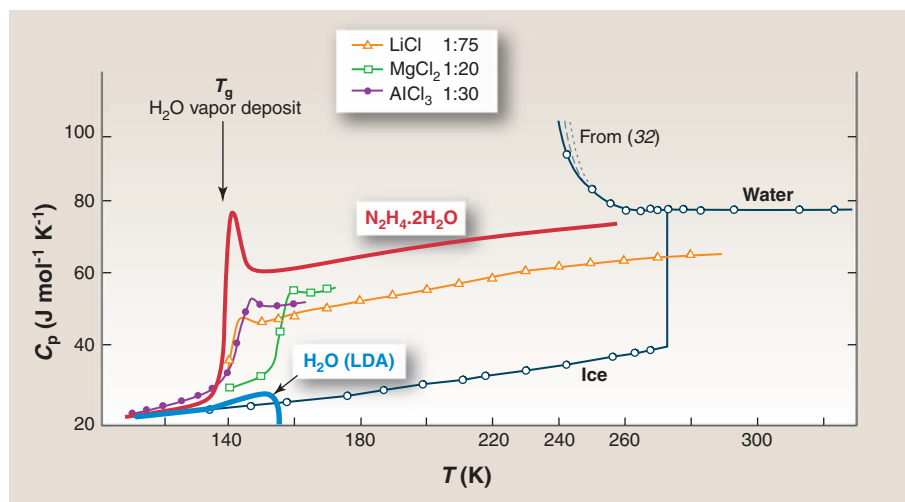
Fig. 1. Heat capacity data on water in supercooled and glassy forms, together with the heat capacities of molecular and ionic solutions in which the majority component is water. In the case of ionic solutions, the heat capacity reported is per mole of water. The plot for LDA is based on (29) (which established an LDA glass heat capacity of 2 J mol^{-1} K^{-1} above ice) and on (20) and (96) (which established the change of slope at 136 K, and crystallization temperature at ~150 K, of annealed hyperquenched water). The curve for supercooled water combines the bulk water values to –30°C (35) with emulsion values to –38°C (32).

Fig. 2. (**A**) The molar excess heat capacity of supercooled water deduced by assuming phase continuity of supercooled water and vitreous ice and requiring adherence of water's properties to the first and second laws of thermodynamics through "no-man's land." The two curves delimit the range of possible C_p variations between 150 and 236 K allowed by the experimental data uncertainties. The shaded area is equivalent to 0.35 kJ. [From (15) by permission of Elsevier] (**B**) The total heat capacity of nanoconfined water, "internal" component, according to the measurements of Maruyama et al. [original data, as plotted in (38, 39)]. The peak temperature agrees well with the thermodynamic curve of upper panel for the more spread-out limit Also, the heat capacity at the peak is near to the sum of this excess C_p in (A) and the heat capacity of ice at 225 K. Note the absence of any normal glass transition between 100 and 200 K (as in Fig. 1) visible on the scale of this figure. Small hysteretic anomalies are in fact detectable on more sensitive scales at 160 and 120 K [see (38, 39)]. [Reproduced by permission of Institute of Physics]

dilute solution between Ca^{2+} cations and NO_3^- anions in relation to temperature in the accessible temperature regime. By observing the ion pair distribution in the glass, they could tell, from the established temperature dependence, the "fictive" temperature T_f at which the equilibrium became "frozen in" during the hyperquench. The T_f they determined was high, between 200 and 230 K, hugely different from T_g determined in any other experiment. However, we can support their conclusion with thermodynamic data (also from Mayer's lab), along with Starr et al.'s thermodynamic analysis of Fig. 2A (15), as follows.

The immediate thermal rescan of a hyperquenched glass releases the frozen-in enthalpy in excess of that in the normal cooling glass transition, and is seen as a broad exothermic peak lying below T_g [see (24)]. This peak has been measured in several studies and has been used to determine T_f (47). In propylene glycol, quenched at the same rate as is used to vitrify water, T_f lies

25 K above the normal T_g (i.e., at 190 K). To show this for water, the shaded area under the broader curve of Fig. 2A is the frozen-in enthalpy of the Hallbrucker-Mayer HQG (48). Measured during rescan, as shown by the cross-hatched area in figure 1 of (49), it amounts to 0.19 kJ/mol (1/30th of the heat of crystallization) plus an extra 0.16 kJ/mol needed to allow for the (unusual) way the enthalpy recovery is cut short by the recrystallization in the case of HQG (48)—in all, 0.35 kJ/mol.

The upper bound of the shaded area in Fig. 2A should give an independent measure of T_f for the Mayer-Hallbrucker HQG. The result is seen in Fig. 2 is $T_f = 200$ K, which falls at the lower end (within the broad uncertainty range) of the ion pair probe method (46). At this temperature, we calculate [using the activation energy for diffusion in glassy water, 45 kJ/mol, reported by Kay at recent meetings (16), and assuming a preexponential constant of 10^{-14} s] that the relaxation time should

be about 10^{-4} s—which accords with the finding for hyperquenched propylene glycol (50).

In summary, there is evidence for a sudden change at around 200 K in the ability of water to equilibrate during hyperquenching. This value lies below the temperature of the C_p peak and presumably is the glass transition displaced to very high temperatures, in accord with Eq. 1, by the combination of high quenching rate and the Arrhenius ("strong liquid") character of the relaxation time in water below its C_p peak. This is fully consistent with the finding (51) that, during ultrafast heating, no glass transition is seen below 205 ± 5 K where crystallization occurs.

Order-Disorder Transitions

The above discussion builds a self-consistent picture, but I now argue that the understanding of water, given the limits on the direct measurements we can make, depends on our ability to integrate this tentative picture with the behavior of related systems and processes that are better understood—in this case, past studies of order-disorder processes (where the term is used in its older, general, pre-universality class, sense, as we are not concerned with behavior near the peak). Of special relevance are those cases that have been shown to have ergodicity-breaking "transitions" in their tail ends.

The most directly comparable case we find is that of the rotator phase of the fullerene C_{60}, which is dielectrically active in its crystalline state. This system has been studied by many workers using different techniques, particularly dielectric relaxation (52), heat capacity (53), and the related enthalpy relaxation process (53). Some data are reproduced in Fig. 3. The total heat capacity exhibits a sharp peak at 250 K—the rotator phase transition (Fig. 3A). The relaxation time for reorientation of the polar defects has been measured by dielectric relaxation (52) and is perfectly Arrhenius in character between 100 K and the phase transition at 250 K (11 orders of magnitude in τ). It blends smoothly with the longer relaxation times measured by enthalpy relaxation, which are shown in the Arrhenius plot of Fig. 3C. The value at 90 K is 1000 s and is characteristic of the glass transition when measured at the slow effective scan rate of adiabatic calorimetry, 0.1 K/min. In differential scanning calorimetry (DSC) at 10 K/min, it would be observed where $\tau_H = 100$ s (i.e., at 91.3 K). This experiment identifies unambiguously the tiny step at 90 K as the ergodicity-restoring glass transition for this system, which is shown on an expanded scale in Fig. 3B. It is thus a "glassy crystal" but one that terminates its configurational excitation in a continuous order-disorder transition (λ transition) rather than in a first-order melting transition, as in the usual case. This happens because fullerenes cannot melt even at high pressures, in consequence of their critical temperatures falling within their solid-phase stability domains (54). They can only sublime.

We see that the magnitude of the glass transition in C_{60} is even smaller than that for water. Being so weak, it would not be detected using ordinary DSC

procedures. The phase transition temperature, 250 K, lies somewhat above that for water. The greater disordering range, 91 to 250 K, is consistent with the smaller activation energy for excitation [28.6 kJ/mol (*52*) versus 45 kJ for water (*16*)].

Although no "hyperquench and scan" studies have been made on the very weak "glass transition" seen in the tail of the C_{60} disordering transition (*53*) (Fig. 3B), it is reasonable to expect to see, in such studies, the same phenomena exhibited by water (*48*). Indeed, this should be seen in any of the familiar λ transitions (NaNO$_3$, etc.) when examined with sufficient precision. More important, they may also be seen in cases such as KNO$_3$ in which the λ transition is interrupted by a first-order transition to the higher-temperature, more disordered, phase. According to current practice, all should be termed "glass transitions" because in each case a residual disorder is frozen in below the "glass temperature." That the disordering occurs within a crystal, not a glass, is not of concern here. Premelting phenomena in many crystals fall into the same class, and their fast quenching and reheating behavior should likewise be examined for "glass transitions" associated with the freezing-in of crystalline disorder. All will be very weak, like that of water. The effect of the freezing-in of defect concentration on the ionic conductivity of crystals is known; it resembles the effect of the glass transition on the conductivity of ion-conducting glasses (*55*). Both resemble the "fragile-to-strong" transition currently much discussed in confined water (*56*), but are in fact quite different because they depend entirely on ergodicity-breaking at the transition.

Which type of glass transition is the more "interesting" is a matter of taste, but it is clear that those occurring in the tails of λ transitions are glass transitions that are stripped of the mystery and the challenge of the glass transition that current theory strives to resolve (*8, 57–66*). There is no Kauzmann paradox demanding resolution, and no hidden first- or second-order transitions to be discussed and contested. As a class, λ transitions have been intensively studied, and their heat capacity forms are quite well understood. It is interesting to

note, however, that two of the current group of theoretical approaches to the glass transition (*59, 67*) obtain considerable success in describing the supercooled liquid heat capacity behavior with excitation thermodynamics mathematically equivalent to that of the popular, although approximate, Bragg-Williams theory for the order-disorder transition, originally invented to describe disordering in alloys such as AuCu (1:1). This is

Fig. 3. The λ transition and glass transition for the rotator phase of C_{60}. (**A**) The full transition, showing the peak of the transition (where the order parameter vanishes) at 250 K. The box at the lower left is enlarged in (**B**) to show the heat capacity anomaly (glass transition) at 83 to 88 K. (**C**) Arrhenius plot of the enthalpy relaxation time for C_{60}, from which the relaxation time at the onset of the glass transition is seen to be 104 s. In a "standard" DSC scan, the onset would be at 91.3 K, which can be predicted from Eq. 1 using the activation energy measured in (*52*). [From (*53*) by permission]

also the property of the Moynihan model of polyamorphism in water (*68*) discussed below.

The Energy Landscape

What distinguishes a λ transition with ergodicity breaking (i.e., glass transition) in its tail, as in Fig. 3, from a "normal" glass transition? Operationally, it is that in the former, the peak in the configurational heat capacity lies in the ergodic domain, whereas in the latter, it is hidden below the glass transition. Kinetically, it reflects the ease with which the system can explore its potential energy hypersurface without getting "stuck." Thus, the potential energy landscape (PEL) for water must have very shallow "basins of attraction" (*69*) so that it can deexcite with temperature decrease, even at low temperatures. On accessible time scales, it can then approach closely the global minimum on its PEL. The glass transition interrupts only the final approach to the global minimum, so that the "ideal" glass state becomes (almost) accessible in the same way (*70*) as the "folding funnel" (*71*) makes the global minimum accessible to protein molecules. The entropy for water near its T_g is indeed remarkable for its small value (*72, 73*), rivaled only by vitreous BeF$_2$ (*42*).

Although the ergodicity-breaking phenomenon in water may therefore be simpler to interpret than that in normal glass-formers, it is by no means less important. Indeed, it allows us to shine light on broader aspects of viscous liquid thermodynamics and on the whole strong/fragile liquid phenomenology of glass-formers, as described in (*42*).

The Fragile-to-Strong Crossover in Supercooled Water

Let us consider how all the foregoing can help us understand another striking feature of the behavior of liquid water, the fragile-to-strong transition, that is observed indirectly upon hyperquenching. Long only an inferred phenomenon (*13–15*), this abrupt change in the temperature dependence of relaxation processes in supercooling water (*15*) has now been directly observed, for both diffusivity and relaxation time, by Chen and colleagues (*56, 74, 75*). They achieved this by study of water samples sequestered in nanoscopic (1.5 nm diameter) tubule arrays, in which the water does not crystallize. The most extensive data come from the pulsed field gradient spin-echo diffusivity study of Mallamace et al. (*56*) in which the break in temperature dependence of the diffusivity occurs at 220 K—the same temperature seen in Fig. 2 for the peak in the nanosample heat capacity (*38, 39*). Although the nano-sequestration clearly changes the behavior of the water (*40, 41, 76*), it is to damp out, rather than to exaggerate, the behavior in real water. These changes are attributed to a Widom (supercritical fluctuations) line crossing by Stanley and co-workers (*77*) (see below). Alternatively, some authors attribute them to confinement length scale effects (*78*). In any case, a strong-fragile transition is a direct prediction of Fig. 2 via the Adam-Gibbs equation (*79*), $\tau = \tau_0 \exp(C/TS_c)$, where S_c is the configurational entropy and τ_0 and C are

constants. This relation requires the temperature dependence of any relaxation process to change sharply when the heat capacity peaks, as demonstrated in figure 3 of (15).

Order-Disorder Versus "Second Critical Point" Scenarios

As already stated, the present interpretation invoking order-disorder phenomenology is compatible with a critical point interpretation of supercooled water phenomenology. However, it does not require that a critical point exist. The majority of order-disorder processes end in first-order phase transitions to the disordered state. For bulk water at ambient pressure and excluding crystallization, it remains possible that the thermal excitation of the low-density liquid (LDL) phase [or the supercooling of the high-density liquid (HDL) phase] may terminate in first-order phase transitions. Indeed, a first-order phase transition is what is implied by the power-law fits of data in much of the supercooled water (HDL) literature (1, 2, 35, 80, 81) [and also what is found (67) in an excitations model best fit to the low-temperature LDL (confined water) heat capacity data in Fig. 2]. It is also what is required by the phase diagram derived (82) from the Haar-Gallagher-Kerr equation of state, which was long considered the best description of water in its stable state, and which has proven capable of predicting the behavior of water in the homogeneous nucleation (kinetic) limit of its metastable stretched state (83). This phase diagram predicts distinct and different tensile limits for HDL and LDL.

Power laws imply diverging fluctuations. They arise by approach either to a critical point or to one of the spinodal lines that radiate out from a critical point. In the latter case, a first-order transition must be crossed before the critical divergence temperature is reached, as seen in the case of the water analog, liquid silicon, in (84).

The properties that have been measured to the lowest temperature in ambient-pressure supercooled water are the spin-lattice relaxation time T_1 and the viscosity. T_1 can be studied in the presence of crystalline ice (e.g., in emulsion samples) without loss of precision because the crystal signal is not motionally narrowed. In (2), data down to 238 K were fit by a power law with the same singular temperature, 228 K (233 K for D_2O), that had earlier been obtained from fits to data on other properties (5, 85, 86). Note that there is only a narrow range, 10 K, in which to encounter either a major deviation from the power law or, alternatively, the intervening liquid-liquid phase transition. Only the T_1 plots for protons (or deuterons) show any deviations from the power law, and protons (deuterons) are known to decouple from oxygen motions by the Grotthus mechanism. It is plausible that the sharp, seemingly inevitable, crystallization that occurs at 234 K in these slowly cooled samples is the result of encountering the phase transition to the fast-crystallizing LDL phase, because this would revisit what has already been seen repeatedly (84, 87)

in liquid silicon, and indeed in TIP4P water (Ewald summation version) in simulation (88).

What is different here from the original Speedy-Angell proposition (85, 86) is the recognition that the spinodal, rather than existing alone as the unique endpoint of the liquid state, in fact underlies a phase transition to the new liquid phase, LDL, as indeed the phase diagram of (82) would imply. Interestingly enough, the first-order transition temperature obtained in (67) from the fit of an order-disorder excitations theory to the confined water LDL C_p data seen in Fig. 2 falls at 229 K (67)—slightly above the HDL spinodal, as it should. Furthermore, this must be a minimum value because it is based on fitting data for water in hydrophilic confinement, for which the critical point is [according to simulations (41)] pushed from negative to positive pressures. According to the phase diagram of (82), this first-order liquid-liquid transition description actually belongs to a "critical point–free" scenario, which is quite distinct from both "second critical point" (3, 31, 43) and "singularity-free" (3, 31, 89) alternatives. In this respect it lies at the opposite extreme from the "singularity-free" scenario (89), which requires T_c to retreat to 0 K (3). The critical point–free scenario, or a variant in which the second critical point exists but only at negative pressure, are supported by the observation that all ambient-temperature response functions and transport coefficients appear to be diverging at the same singular temperature (2, 80, 81, 85, 86), a behavior not found much beyond the critical point (77).

Countering the critical point–free scenario are two cooperative excitation models, the Moynihan two species–nonideal solution model (68) and the Franzese-Marques-Stanley intramolecular coupling model (90, 91), each of which yields a liquid-liquid critical point at positive pressure. In this case the fragile-strong transition becomes, at ambient pressure, a continuous transition associated with crossing the "Widom line" (the higher-order extension of the coexistence line beyond the critical point). The Moynihan model in particular inputs data from the experimental hysteresis of the LDA-HDA polyamorphic transition (27) as spinodal pressure limits, and derives a critical pressure of 380 atm. This is remarkably close to the value of 270 atm deduced by Fuentavilla and Anisimov (92) from critical scaling law fitting of the bulk water heat capacity and compressibility data. The latter analysis gave a critical temperature of 232 K, which is unfeasibly high, whereas the Moynihan model gives 225 K. However, the absence of any sign of power-law breakdown for the ambient-pressure properties, except for those known to decouple, is a problem for these models. The ^{17}O T_1 measurements in coldest emulsions (93) bear repeating with modern equipment because they appear to be crucial in determining the most plausible description of this important liquid.

Concluding Remarks

There is not much distinction between the order-disorder transition scenario we have

outlined here and the second critical point scenario, which associates all water anomalies with the existence of a second critical point. Is the second critical point to be regarded as the source of the anomalies of water, or is the cooperativity of the configurational excitations (implied by the form of the heat capacity that we have extracted) to be seen as the primary phenomenon to be interpreted—one that may, at some parameter or some thermodynamic field choice, produce a critical point? It becomes a "chicken or egg" issue, and the question is: Which is scientifically the more fruitful? When a critical point exists, all of the universalities that are implied (near the transition) can be enjoyed, but if the cooperativity is insufficient to induce a critical point, the liquid can still have interesting and important practical properties, such as a vanishing expansion coefficient (as in SiO_2). Indeed, the relation of water to SiO_2 glass and liquid, and the "big picture" of glass-formers, is a vitally interesting question—which is taken up elsewhere (42).

References and Notes

1. E. W. Lang, H. D. Lüdemann, Angew. Chem. Int. Ed. Engl. 21, 315 (1982).
2. C. A. Angell, Annu. Rev. Phys. Chem. 34, 593 (1983).
3. P. G. Debenedetti, J. Phys. Condens. Matter 15, R1669 (2003).
4. A "phase transition" is a "singularity," an event that totally separates the substance on one side of the transition from the substance on the other side. For a pure substance, this happens in a reversible manner. For instance, in the case of melting at constant pressure, once the pressure is fixed there is only one temperature at which the two phases, liquid and crystal, can coexist. At any other temperature, it is one or the other of the two—an "all-or-nothing" situation.
5. M. Oguni, C. A. Angell, J. Chem. Phys. 73, 1948 (1980).
6. We like to define T_g as the onset value obtained when scanning at 20 K/min, after cooling at a standard rate of 20 K/min, because at that temperature the relaxation time is close to 100 s and therefore agrees with the value reported by relaxation spectroscopists, who define T_g as the temperature at which the relaxation time is 100 s (94).
7. F. Stickel, E. W. Fischer, R. Richert, J. Chem. Phys. 102, 6251 (1995).
8. J. Dyre, Rev. Mod. Phys. 78, 953 (2006).
9. C. A. Angell, J. Therm. Anal. Calorim. 69, 785 (2002).
10. C. A. Angell, Science 267, 1924 (1995).
11. N. Giovambattista, C. A. Angell, F. Sciortino, H. E. Stanley, Phys. Rev. Lett. 93, 047801 (2004).
12. A. Minoguchi, R. Richert, C. A. Angell, Phys. Rev. Lett. 93, 215703 (2004).
13. C. A. Angell, J. Phys. Chem. 97, 6339 (1993).
14. K. Ito, C. T. Moynihan, C. A. Angell, Nature 398, 492 (1999).
15. F. Starr, C. A. Angell, H. E. Stanley, Physica A 323, 51 (2003).
16. B. D. Kay et al., Abstracts of papers of the American Chemical Society 229: U706-U706 024-PHYS Part 2 (2005).
17. S. M. McClure, D. J. Safarik, T. M. Truskett, C. B. Mullins, J. Phys. Chem. B 110, 11033 (2006).
18. D. Kivelson, G. Tarjus, J. Phys. Chem. B 105, 6220 (2001).
19. R. S. Smith, B. D. Kay, Nature 398, 788 (1999).
20. I. Kohl, L. Bachmann, E. Mayer, A. Hallbrucker, T. Loerting, Nature 435, E1 (2005).
21. The most recent assessment of the laboratory glass transition in water (20) reports $\Delta C_p = 0.7$ J K^{-1} mol^{-1} compared with 35 J K^{-1} mol^{-1} for $(H_2O)_2 \cdot H_2O_2$.
22. J. A. Pryde, G. O. Jones, Nature 170, 685 (1952).
23. D. R. MacFarlane, C. A. Angell, J. Phys. Chem. 88, 759 (1984).
24. Y. Z. Yue, C. A. Angell, Nature 427, 717 (2004).
25. E. Mayer, J. Appl. Phys. 58, 663 (1985).
26. E. F. Burton, W. F. N. Oliver, Nature 135, 505 (1935).

27. O. Mishima, *J. Chem. Phys.* **100**, 5910 (1994).
28. A highly analogous behavior has recently been observed for the element germanium, which shares with water an open tetrahedral crystal ground state and an increase in density upon fusion. It can be obtained in the amorphous form by different methods, vapor deposition, electrodeposition, and quenching of flame-formed droplets, and these forms are apparently identical. It has recently been shown (*95*) that a dense glassy form, obtained by quenching the high-pressure metallic liquid, undergoes a polyamorphic phase change, during pressure decrease at ambient temperature, to yield a glass that is essentially the same as that formed at low pressure, although with small displacements in the structure factor $S(q)$ peak positions.
29. Y. P. Handa, D. D. Klug, *J. Phys. Chem.* **92**, 3323 (1988).
30. A. Hallbrucker, E. Meyer, G. P. Johari, *J. Phys. Chem.* **93**, 4986 (1989).
31. H. E. Stanley, P. G. Debenedetti, *Phys. Today* (2005).
32. C. A. Angell, W. J. Sichina, M. Oguni, *J. Phys. Chem.* **86**, 998 (1982).
33. D. H. Rasmussen, A. P. Mackenzie, J. C. Tucker, C. A. Angell, *Science* **181**, 4079 (1973).
34. G. P. Johari, *J. Chem. Phys.* **107**, 10154 (1997).
35. E. Tombari, C. Ferrari, G. Salvetti, *Chem. Phys. Lett.* **300**, 749 (1999).
36. O. Mishima, *J. Chem. Phys.* **121**, 3161 (2004).
37. Y. J. Kim, B. E. Wyslouzil, G. Wilemski, J. Wölk, R. Strey, *J. Phys. Chem. A* **108**, 4365 (2004).
38. S. Maruyama, K. Wakabayashi, M. A. Oguni, *AIP Conf. Proc.* **708**, 675 (2004).
39. M. Oguni, S. Maruyama, K. Wakabayashi, A. Nagoe, *Chem. Asian* **2**, 514 (2007).
40. C. A. Angell, *Nature Nanotechnol.* **2**, 396 (2007).
41. I. Brovchenko, A. Oleinikova, *J. Chem. Phys.* **126**, AN214701 (2007).
42. C. A. Angell, *Bull. Mater. Res. Soc.* (Turnbull lecture) (February 2007); arxiv.org/abs/0712.4233.
43. P. H. Poole, F. Sciortino, U. Essmann, H. E. Stanley, *Nature* **360**, 324 (1992).
44. F. Sciortino, P. H. Poole, U. Essmann, H. E. Stanley, *Phys. Rev. E* **55**, 727 (1997).
45. H. E. Stanley *et al.*, *Physica A* **236**, 19 (1997).
46. G. Fleissner, A. Hallbrucker, E. Mayer, *J. Phys. Chem. B* **102**, 6239 (1998).

47. L.-M. Wang, C. A. Angell, *J. Non-Cryst. Solids* **353**, 3829 (2007).
48. A. Hallbrucker, E. Mayer, G. P. Johari, *J. Phys. Chem.* **93**, 4986 (1989).
49. V. Velikov, S. Borick, C. A. Angell, *Science* **294**, 2335 (2001).
50. C. A. Angell *et al.*, *J. Phys. Condens. Matter* **15**, S1051 (2003).
51. M. Chonde, M. Brindza, V. Sadtchenko, *J. Chem. Phys.* **125**, 094501 (2006).
52. P. Mondal, P. Lunkenheimer, A. Loidl, *Z. Phys. B* **99**, 527 (1996).
53. T. Matsuo *et al.*, *Solid State Commun.* **83**, 711 (1992).
54. M. H. J. Hagen, E. J. Meijer, G. C. A. M. Moolj, *Nature* **365**, 425 (1993).
55. F. Mizuno *et al.*, *J. Non-Cryst. Solids* **352**, 5147 (2006).
56. F. Mallamace *et al.*, *J. Chem. Phys.* **124**, 161102 (2006).
57. H. Tanaka, *J. Phys. Cond. Matter* **111**, 3175 (1999).
58. H. Tanaka, *Phys. Rev. E* **62**, 6978 (2000).
59. C. A. Angell, C. T. Moynihan, *Metall. Mater. Trans. B* **31**, 587 (2000).
60. X. Xia, P. G. Wolynes, *Proc. Natl. Acad. Sci. U.S.A.* **97**, 2990 (2000).
61. D. V. Matyushov, C. A. Angell, *J. Chem. Phys.* **123**, 034506 (2005).
62. H. Tanaka, *Phys. Rev. Lett.* **90**, 055701 (2003).
63. J. P. Garrahan, D. Chandler, *Proc. Natl. Acad. Sci. U.S.A.* **100**, 9710 (2003).
64. G. Biroli, J.-P. Bouchaud, G. Tarjus, *J. Chem. Phys.* **123**, 044510 (2005).
65. D. Chandler, J. P. Garrahan, *J. Chem. Phys.* **123**, 044511 (2005).
66. P. G. Debenedetti, F. H. Stillinger, M. S. Shell, *J. Phys. Chem. B* **107**, 14434 (2003).
67. D. V. Matyushov, C. A. Angell, *J. Chem. Phys.* **126**, 094501 (2007).
68. C. T. Moynihan, *Proc. Mat. Res. Soc. Symp.* **455**, 411 (1997).
69. F. H. Stillinger, *Science* **267**, 1935 (1995).
70. C. A. Angell, *Philos. Trans. R. Soc. London Ser. A* **363**, 415 (2005).
71. P. G. Wolynes, J. N. Onuchic, D. Thirumalai, *Science* **268**, 960 (1995).
72. R. J. Speedy, P. G. Debenedetti, R. S. Smith, C. Huang, B. D. Kay, *J. Chem. Phys.* **105**, 240 (1996).
73. E. Whalley, D. D. Klug, Y. P. Handa, *Nature* **342**, 782 (1989).
74. A. Faraone, L. Liu, C. Y. Mou, C. W. Yen, S. H. Chen, *J. Chem. Phys.* **121**, 10843 (2004).

75. L. Liu, A. Faraone, C. Mou, C. W. Yen, S. H. Chen, *J. Phys. Condens. Matter* **16**, S5403 (2004).
76. J. Rault, R. Neflati, P. Judenstein, *Eur. Phys. J. B* **36**, 627 (2003).
77. L.-M. Xu *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **102**, 16558 (2005).
78. J. Swenson, H. Jansson, R. Bergman, *Phys. Rev. Lett.* **96**, 247802 (2006).
79. G. Adam, J. H. Gibbs, *J. Chem. Phys.* **43**, 139 (1965).
80. B. D. Cornish, R. J. Speedy, *J. Phys. Chem.* **88**, 1888 (1984).
81. W. S. Price, H. Ide, Y. Arata, *J. Phys. Chem. A* **103**, 448 (1999).
82. Q. Zheng *et al.*, in *Liquids Under Negative Pressures*, A. R. Imre, H. J. Maris, P. R. Williams, Eds. (Kluwer Academic, Dordrecht, Netherlands, 2002), pp. 33–46.
83. Q. Zheng, D. J. Durben, G. H. Wolf, C. A. Angell, *Science* **254**, 829 (1991).
84. S. Sastry, C. A. Angell, *Nat. Mater.* **2**, 739 (2003).
85. R. J. Speedy, C. A. Angell, *J. Chem. Phys.* **65**, 851 (1976).
86. R. J. Speedy, *J. Phys. Chem.* **86**, 982 (1982).
87. C. A. Angell, S. Borick, M. Grabow, *J. Non-Cryst. Solids* **205–207**, 463 (1996).
88. M. Matsumoto, A. Saito, I. Ohmine, *Nature* **416**, 409 (2002).
89. S. Sastry, P. G. Debenedetti, F. Sciortino, H. E. Stanley, *Phys. Rev. E* **53**, 6144 (1996).
90. G. Franzese, M. I. Marques, H. E. Stanley, *Phys. Rev. E* **67**, 011103 (2003).
91. G. Franzese, H. E. Stanley, *J. Phys. Cond. Matter* **19**, 205126 (2007).
92. D. A. Fuentevilla, M. A. Anisimov, *Phys. Rev. Lett.* **97**, 195702 (2006).
93. J. C. Hindman, *J. Chem. Phys.* **60**, 4488 (1974).
94. C. A. Angell, *Chem. Rev.* **102**, 2627 (2002).
95. H. Bhat *et al.*, *Nature* **448**, 787 (2007).
96. K. Hofer, E. Mayer, G. P. Johari, *J. Phys. Chem.* **95**, 7100 (1991).
97. Supported by the NSF under Solid State Sciences grant DMR 0454672 and Chemical Sciences grant 0404714. I thank V. Molinero and P. G. Debenedetti for important critical comments on earlier versions of this manuscript, and H. E. Stanley, S. Buldyrev, F. Sciortino, C. T. Moynihan, and D. V. Matyushov for helpful discussions.

Languages Evolve in Punctuational Bursts

Quentin D. Atkinson,[1]* Andrew Meade,[1] Chris Venditti,[1] Simon J. Greenhill,[2] Mark Pagel[1,3]†

American English emerged abruptly when Noah Webster introduced his *American Dictionary of the English Language*, insisting that "as an independent nation, our honor requires us to have a system of our own, in language as well as government" (*1*).

descendant language. If lexical divergence is a gradual process that is not affected by the emergence of a new language, then the path length or total distance from the root of the tree to the languages at the tips should be independent of the number of language-splitting events or nodes

among these languages (Fig. 1C, purple bar). These effects are comparable in size to punctuational genetic changes observed among biological species (~22%; Fig. 1C, yellow bar) (*5*).

Our results, representing thousands of years of language evolution, identify a general tendency for newly formed sister languages to diverge in their fundamental vocabulary initially at a rapid pace, followed by longer periods of slower and gradual divergence. Punctuational bursts in phonology, morphology, and syntax, or at later times of language contact, may also occur. Linguistic founder effects could cause these rapid changes if newly formed languages emerge in small groups, such as in Austronesian. Alternatively, as the example of American English illustrates, speakers often use language not just as a means of communication but as a tool with social functions, including promoting cohesion and group identity (*6*, *7*). Punctuational language change may thus reflect a human capacity to rapidly adjust languages at critical times of cultural evolution, such as during the emergence of new and rival groups.

Fig. 1. Inferring punctuational language evolution. (**A**) Tree of four languages. If language-splitting events (red nodes) cause bursts of change, the paths from the root to a and b should be longest, followed by c then d (*8*); here, they are all equal. (**B**) Root-to-tip path length plotted against number of nodes along each path for punctuational trees in Bantu (orange), Indo-European (blue), Austronesian (green), and Polynesian (purple). Fitted lines show the relationship between path length and nodes after controlling for phylogeny (*8*). A positive slope is indicative of punctuational evolution. Path lengths for each data set were scaled to account for the number of characters examined. (**C**) Histogram showing the percentage of lexical evolution attributable to punctuational bursts at language-splitting events (mean ± SD) for Bantu (B, orange), Indo-European (IE, blue), Austronesian (A, green), and Polynesian (P, purple) (*8*). For comparison, the percentage of molecular evolution attributable to punctuational effects in biological species is also shown (S, yellow) (*4*).

Punctuational or rapid bursts of change associated with the emergence of new languages, or at later language contact, have been proposed as an important feature of language evolution (*2*, *3*), echoing a long-standing parallel debate in biology (*4*, *5*). Webster's actions illustrate how the desire for a distinct social identity may cause languages to change rapidly (*6*, *7*), but whether such punctuational change is a regular feature of human language evolution has never been demonstrated (*2*). With use of three of the world's major language families comprising over one-third of all the world's languages, we found that punctuational bursts of change at the time of language splitting are an important and general process in language evolution and account for 10 to 33% of the total divergence among these languages in their fundamental vocabularies.

We studied punctuational evolution in phylogenetic trees of language families inferred from vocabulary data (*8*). These trees describe the separate paths of evolution leading from a common ancestral language to the set of observed extant languages at the tips of the tree (Fig. 1A). The lengths of the individual branches of the trees record the amount of lexical divergence (replacement of words) between an ancestral and a

found along that path. If language-splitting events produce punctuational bursts of evolution, however, we expect to find more total lexical divergence (longer path lengths) along paths through the tree that record more language-splitting events (*4*, *8*, *9*).

In each language family, we found significantly more lexical change along paths in which more new languages have emerged, the signature of punctuational evolution (Fig. 1B). These results take into account the phylogenetic relationships among languages, control for a well-known artifact of phylogenetic reconstruction (*10*), and cannot be attributed to borrowing of vocabulary (*8*). The punctuational effects account for a surprising amount of the total lexical divergence among the languages (Fig. 1C): 31% of vocabulary differences among Bantu language speakers arose at or around the time of language-splitting events, 21% among Indo-European languages, and 9.5% in Austronesian (*8*). In the settlement of the Pacific, successive founder events (a small group colonizing a new location) by Polynesian language speakers may have caused increased rates of language change (*11*). Consistent with this, we inferred a stronger punctuational effect in the Polynesian subclade of the Austronesian tree, contributing to about 33% of lexical differences

References and Notes

1. N. Webster, *Dissertations on the English Language* (Isaiah Thomas, Boston, 1789), p. 20.
2. R. D. Janda, B. D. Joseph, in *The Handbook of Historical Linguistics*, B. D. Joseph, R. D. Janda, Eds. (Blackwell, Oxford, 2003), pp. 3–180.
3. R. M. W. Dixon, *The Rise and Fall of Languages* (Cambridge Univ. Press, Cambridge, 1997).
4. N. Eldredge, S. J. Gould, in *Models in Paleobiology*, T. J. M. Schopf, Ed. (Freeman, San Francisco, 1972), pp. 82–115.
5. M. Pagel, C. Venditti, A. Meade, *Science* **314**, 119 (2006).
6. W. Labov, *Principles of Linguistic Change: Internal Factors* (Blackwell, Oxford, 1994).
7. J. K. Chambers, *Sociolinguistic Theory: Linguistic Variation and Its Social Significance* (Blackwell, Cambridge, MA, 1995).
8. Materials and methods are available on *Science* Online.
9. A. J. Webster, R. J. Payne, M. Pagel, *Science* **301**, 478 (2003).
10. C. Venditti, A. Meade, M. Pagel, *Syst. Biol.* **55**, 637 (2006).
11. P. V. Kirch, R. C. Green, *Curr. Anthropol.* **28**, 431 (1987).
12. Supported by a Leverhulme Trust grant to M.P.

[1]School of Biological Sciences, University of Reading, Reading RG6 6AS, UK. [2]Department of Psychology, University of Auckland, Private Bag 92019, Auckland 1142, New Zealand. [3]Santa Fe Institute, 1399 Hyde Park Road, Santa Fe, NM 87501, USA.

*Present address: Institute of Cognitive and Evolutionary Anthropology, University of Oxford, Oxford OX2 6QS, UK.
†To whom correspondence should be addressed. E-mail: m.pagel@reading.ac.uk

Widespread Genetic Incompatibility in *C. Elegans* Maintained by Balancing Selection

Hannah S. Seidel,*† Matthew V. Rockman,*† Leonid Kruglyak*

Natural selection is expected to eliminate genetic incompatibilities from interbreeding populations. We have discovered a globally distributed incompatibility in the primarily selfing species *Caenorhabditis elegans* that has been maintained despite its negative consequences for fitness. Embryos homozygous for a naturally occurring deletion of the zygotically acting gene *zeel-1* arrest if their sperm parent carries an incompatible allele of a second, paternal-effect locus, *peel-1*. The two interacting loci are tightly linked, with incompatible alleles occurring in linkage disequilibrium in two common haplotypes. These haplotypes exhibit elevated sequence divergence, and population genetic analyses of this region indicate that natural selection is preserving both haplotypes in the population. Our data suggest that long-term maintenance of a balanced polymorphism has permitted the incompatibility to persist despite gene flow across the rest of the genome.

*C*aenorhabditis elegans is a globally distributed species of free-living bacteria-eating nematode. Although rare males contribute at a low rate to outcrossing, *C. elegans* occurs primarily as inbred self-fertilizing hermaphrodites (*1–4*). A wild isolate from Hawaii, CB4856, has been identified among well-studied isolates as the most divergent at the sequence level from the standard laboratory strain, N2, derived from an isolate from Bristol, England (*5–7*). As a result of this sequence divergence, the Hawaiian strain is widely used to map mutations induced in the Bristol background.

Genetic incompatibility between Bristol and Hawaii. We generated recombinant inbred lines (RILs) from the 10th generation of an advanced intercross between Bristol and Hawaii to

Lewis-Sigler Institute for Integrative Genomics and Department of Ecology and Evolutionary Biology, Princeton University, Princeton, NJ 08544, USA.

*To whom correspondence should be addressed. E-mail: hseidel@princeton.edu (H.S.S.); mrockman@princeton.edu (M.V.R.); leonid@genomics.princeton.edu (L.K.)
†These authors contributed equally to this work.

study natural genetic variation in *C. elegans*, and we genotyped the RILs at 1450 single-nucleotide polymorphism (SNP) markers (*8*). We noted that a region on the left arm of chromsome I exhibited a dramatic deficit of Hawaii alleles among the RILs. Of 239 RILs, only 5 carried the Hawaii allele at the most skewed marker, and simulations of the intercross pedigree indicated that this allele frequency skew could not have arisen by drift, suggesting that selection had acted during construction of the RILs (fig. S1). We then crossed Hawaii to a Bristol strain carrying a visible marker located 10 cM from the most skewed RIL marker, and we examined F_2 progeny produced by self-fertilizing F_1 hermaphrodites. Surprisingly, approximately 25% of F_2 progeny arrested as embryos, and embryonic lethality segregated opposite the visible marker (Table 1). F_2 lethality was not an effect of the marker: Self-fertilizing F_1 hermaphrodites derived from reciprocal crosses between Hawaii and wild-type Bristol produced 25% dead embryos, as did F_1 hermaphrodites mated to F_1 males (Fig. 1).

Lethality caused by a paternal effect by zygotic interaction. The segregation of embryonic lethality opposite the visible marker implied that the arrested embryos represented those homozygous for the Hawaii allele of a locus linked to the marker, on the left arm of chromosome I. Because embryonic lethality within the Hawaii strain itself is very low (less than 1%), we reasoned that F_2 lethality reflected an incompatibility between the Hawaii allele of this locus and an element in the Bristol genome. We also reasoned that it did not reflect two synthetically lethal alleles segregating in the F_2 population because such an interaction would affect less than one-quarter of F_2 embryos (up to 3/16, depending on linkage and dominance). One-quarter lethality is expected, however, if the incompatibility involves an interaction between the genotype of the zygote and a maternal or paternal effect.

To test this possibility, Hawaii × Bristol F_1 males and hermaphrodites were separately backcrossed to Hawaii individuals, and lethality was scored among the resulting embryos. We observed 50% lethality when F_1 males were mated to Hawaii hermaphrodites but less than 2% lethality in the reciprocal cross (Fig. 1). Thus, lethality depends on both paternal and zygotic genotype, but is independent of maternal cytoplasm. (Both Hawaii and F_1 hermaphrodites produced dead embryos, 50% and 25%, respectively, when mated to F_1 males.) In sum, lethality appears to result from a paternal effect by zygotic interaction, whereby embryos homozygous for the Hawaii allele of a zygotically acting locus fail to hatch when the sperm parent—male or hermaphrodite—is a Hawaii × Bristol heterozygote. An interaction between a paternal effect and a zygotically acting gene is surprising because sperm-supplied factors are expected to act during fertilization and first cleavage (*9*), whereas early embryogenesis is primarily controlled by maternally contributed factors, and zygotic transcription is not known to occur before the four-cell stage (*10*).

Tight linkage of the zygotically acting and paternal-effect loci. To understand the genetic basis of the incompatibility, we used the RILs

Table 1. F_2 lethality segregates opposite visible marker *bli-3*, located 10 cM from the most skewed RIL marker. Genotypes of F_2 progeny from selfing Bristol/*bli-3* and Hawaii/*bli-3* hermaphrodites were scored. Embryonic lethality from selfing Hawaii/*bli-3* hermaphrodites was slightly greater than 25% because the visible marker introduces a small percentage of lethality.

F_2 genotype	Bristol/*bli-3* hermaphrodite	Hawaii/*bli-3* hermaphrodite
bli-3/*bli-3*	21.7% (128)	21.8% (128)
bli-3/+	49.9% (295)	42.0% (246)
+/+	24.5% (145)	5.1% (30)
Arrested embryos	3.9% (23)	31.1% (182)
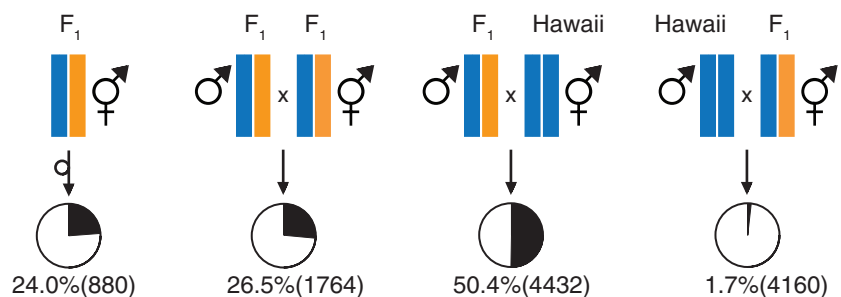
Fig. 1. Paternal effect by zygotic lethality. The percent of embryonic lethality (total) was scored from the crosses shown. Orange and blue indicate Bristol and Hawaii haplotypes, respectively. Pie charts show the proportions of embryos that hatched (white) and failed to hatch (black). F_1 individuals were derived from reciprocal Bristol × Hawaii crosses. Embryonic lethality from selfing Bristol and Hawaii hermaphrodites, reciprocal Bristol × Hawaii, and reciprocal Bristol × F_1 crosses was less than 0.8% (*n* > 240 embryos for each).
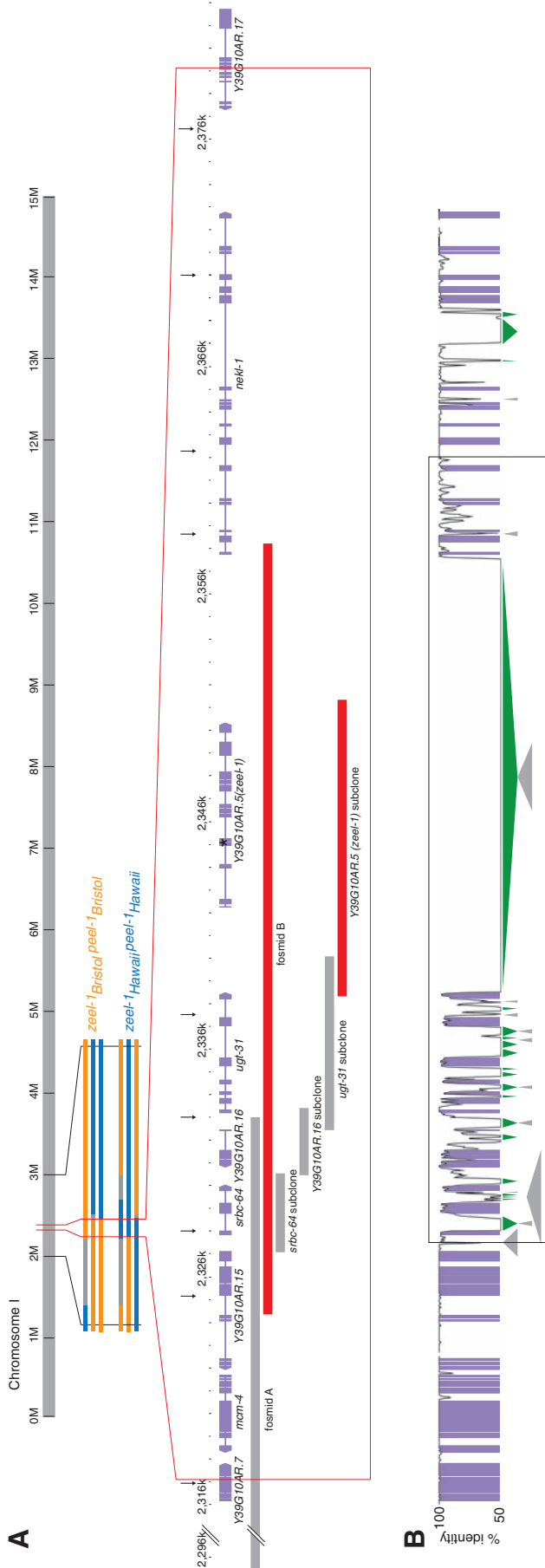
Fig. 2. *zeel-1* and *peel-1* map to the same 62-kb interval. (**A**) Colored bars represent Bristol (orange) and Hawaii (blue) haplotypes carried by six RILs used to map *zeel-1* and *peel-1*. Gray bars show regions between undetermined RIL breakpoints. RIL genotypes inferred from mapping crosses are shown to the right of the haplotypes. The red box in the middle indicates the interval to which *zeel-1* and *peel-1* map. Gene annotations derive from WormBase (ws170) gene models for Bristol sequence. Horizontal bars indicate coverage of fosmid A (WRM0614dC06), fosmid B (WRM0633bE09), and subclones used in transgenic complementation, with rescuing transgenes in red. A black asterisk indicates the frame shift introduced into the *Y39G10AR.5* subclone. Arrows indicate markers used to genotype wild isolates. All markers are in complete linkage disequilibrium. (**B**) mVista alignment of Hawaii and Bristol sequence across a portion of the *zeel-1/peel-1* interval (34). We used default parameters and Bristol as the reference sequence. The black box surrounds the interval of elevated divergence between Bristol and Hawaii. Large (>50 bp) deletions (green) and insertions (gray) in Hawaii relative to the Bristol sequence are shown below the alignment.

to map both the zygotically acting locus, *zeel-1* (*zygotic epistatic embryonic lethal–1*), and the paternal-effect locus, *peel-1* (*paternal effect epistatic embryonic lethal–1*). We crossed RILs to Bristol and Hawaii and scored lethality among embryos laid by self-fertilizing F₁ hermaphrodites and Hawaii hermaphrodites mated to F₁ males. The pattern of lethality among F₂ and backcross embryos was consistent with each RIL carrying either the Bristol alleles of both *zeel-1* and *peel-1* or the Hawaii alleles of both. We identified only one genomic interval in which all lines of the former class carried the Bristol haplotype and all lines of the latter class carried the Hawaii haplotype. Thus, both *zeel-1* and *peel-1* map to this interval, a 62-kb region on chromosome I (position 2,317,234 to 2,379,249) (Fig. 2A). We confirmed tight linkage between the two loci; they do not segregate independently among backcross progeny (table S1).

Incomplete penetrance of the incompatibility. The penetrance of the incompatibility (i.e., the extent of lethality among *zeel-1*$_{Hawaii}$ homozygotes sired by *peel-1* heterozygotes) was complete when oocytes were fertilized by male sperm but incomplete when they were fertilized by hermaphrodite sperm. We collected embryos from self-fertilizing F₁ hermaphrodites and from F₁ hermaphrodites mated to F₁ males and genotyped surviving progeny at the *zeel-1* locus. Among self-progeny, approximately 10% of *zeel-1*$_{Hawaii}$ homozygotes survived to hatching, although most had retarded development and abnormal morphologies (*8*). In contrast, none survived when fertilized by F₁ males. Penetrance of the incompatibility also appeared complete among embryos from F₁ males backcrossed to Hawaii hermaphrodites, as these broods lacked the deformed larvae characteristic of surviving *zeel-1*$_{Hawaii}$ homozygotes.

The morphological defects of surviving *zeel-1*$_{Hawaii}$ homozygotes were highly variable and often similar to the terminal phenotype observed in arrested embryos, which usually showed tissue differentiation but no elongation past the twofold stage. Nevertheless, some *zeel-1*$_{Hawaii}$ homozygotes matured to adulthood and produced progeny. These progeny were entirely wild type, implying that the paternal effect is not caused by heritable defects such as DNA damage or aneuploidy in *zeel-1*$_{Hawaii}$ sperm.

Globally distributed incompatibility. To determine the distribution of alleles causing the Bristol-Hawaii incompatibility in the global *C. elegans* population, we phenotyped 62 wild isolates from 40 localities. From each locality, we phenotyped only strains known to be genetically distinct. We crossed each strain to Bristol and Hawaii and scored lethality among embryos from self-fertilizing F₁ hermaphrodites, F₁ males backcrossed to Hawaii hermaphrodites, and F₁ males backcrossed to hermaphrodites of the wild isolate itself. All but one wild isolate produced a pattern of lethality consistent with carrying either the Bristol alleles of both

peel-1 and *zeel-1* (Bristol-compatible strains) or the Hawaii alleles of both (Hawaii-compatible strains). The exception, a strain collected from Roxel, Germany, was compatible with both Bristol and Hawaii, showing no lethality in crosses to either strain. The global distribution of Bristol-compatible and Hawaii-compatible strains demonstrates that the two classes are not geographically isolated (Fig. 3), which is consistent with an absence of large-scale population structure in *C. elegans* (*1, 5, 11–13*). Both classes were found in many localities, and individual samples of compost from two localities in northern Germany contained both Bristol- and Hawaii-compatible strains, indicating that the two classes co-occur at the most local level (*11*).

Molecular signatures of balancing selection. The interval to which *zeel-1* and *peel-1* map contains a region of dramatically elevated sequence divergence between the Bristol and Hawaii haplotypes. This region spans 33 kb of Bristol sequence and includes four full genes and part of a fifth (Fig. 2B). The Hawaii haplotype contains a 19-kb deletion spanning the gene *Y39G10AR.5*. Divergence within coding segments of the remaining genes averages 5%, which is 50 times higher than previous genome-wide estimates of pairwise divergence from both coding and noncoding sequence (*6, 7*). Non-coding segments in this region are largely un-alignable and contain many insertions and deletions, mainly composed of repetitive elements. The left boundary of the divergent interval is abrupt and is marked by a 1-kb insertion in Hawaii. Genomic divergence within the 13 kb immediately outside the insertion is 0.1%. The right boundary is less abrupt, with divergence falling gradually to 0.7% across 4 kb.

We genotyped the wild isolates with markers located throughout the interval and found that all Hawaii-compatible strains carry Hawaii-like haplotypes, whereas all Bristol-compatible strains carry Bristol-like haplotypes (table S2). The doubly compatible strain carries a Bristol-like haplotype. Linkage disequilibrium among markers within the divergent interval is complete but breaks down for markers 165 kb to the left and 78 kb to the right of the interval.

To understand the cause of elevated polymorphism in the *zeel-1/peel-1* interval, we sequenced the exons and adjacent regions of *srbc-64*, a gene located within the divergent interval, from 45 genetically distinct wild isolates. Elevated polymorphism may result from an elevated mutation rate, an ancient coalescence of neutral alleles, or long-term balancing selection maintaining divergent haplotypes against loss by genetic drift, thereby permitting them to accumulate more mutations than expected under neutrality (*14*). Under balancing selection, most mutations will differentiate the two major haplotype classes, creating an excess of intermediate-frequency alleles. Among the *srbc-64* sequences, we observed 80 polymorphic sites but only six distinct haplotypes, far fewer than expected for a

neutral sample (*P* < 0.0001). Furthermore, the allele frequency spectrum was strongly skewed toward intermediate-frequency alleles (Tajima's $D = 3.56$, $P < 0.0001$). These data are consistent with long-term maintenance by balancing selection.

The *srbc-64* sequences form two distinct, deeply branching clades, reflecting the Bristol-
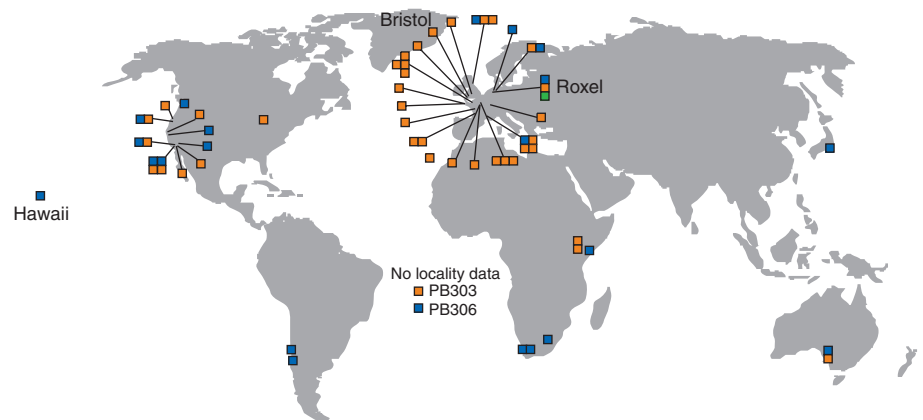
Fig. 3. Strains compatible with Bristol and incompatible with Hawaii (orange) and strains compatible with Hawaii and incompatible with Bristol (blue) are globally distributed. Strains compatible with both Bristol and Hawaii (green) derive from a locality in Roxel, Germany. For each locality, only strains known to be genetically distinct are shown. Strain names and localities are presented in table S7.
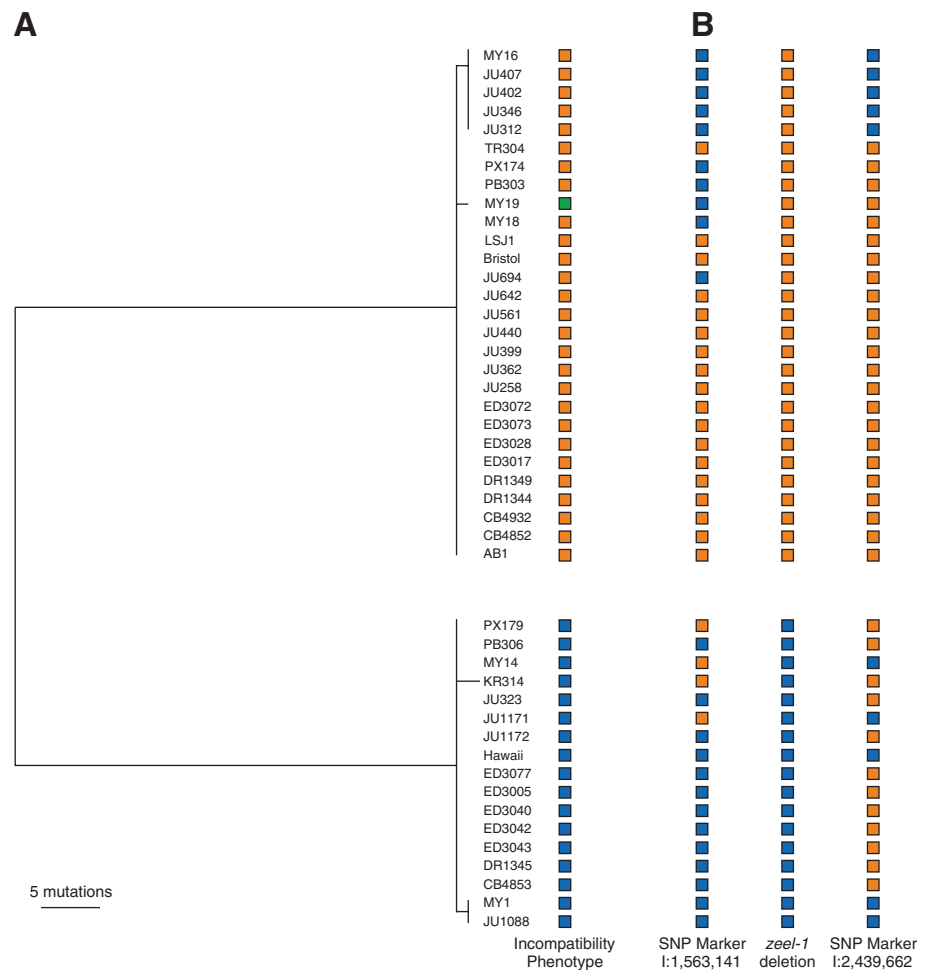
Fig. 4. Signature of balancing selection on the *zeel-1/peel-1* interval. (**A**) Haplotypes of 1193 bp of the *srbc-64* gene, excluding gapped sites, are split into two deeply divergent clades, one compatible with Bristol (orange) and one with Hawaii (blue). Doubly compatible MY19 (green) has a haplotype similar to that of Bristol. (**B**) Recombination has separated the *zeel-1* deletion polymorphism from marker SNPs on both sides of the divergent interval.

and Hawaii-compatible classes (Fig. 4A). This topology is not representative of the genome as a whole. Markers on both sides of the divergent interval exhibit evidence of genetic exchange between the two classes (Fig. 4B), and phylogenies constructed from mitochondrial sequence and nuclear sequences located elsewhere in the genome do not resolve Bristol- and Hawaii-compatible strains into two distinct clades (5, 15), indicating that gene flow occurs across the rest of the genome.

The exceptional divergence between the Bristol and Hawaii haplotypes does not appear to be due to diversifying selection favoring amino acid–changing substitutions, nor to relaxed selection allowing degeneration of their protein-coding sequences. Despite synonymous divergence orders of magnitude above that observed for genes outside the interval, genes within the interval exhibit ratios of nonsynonymous to synonymous substitution of less than 0.2, consistent with purifying selection (Table 2). Given a direct estimate of the mutation rate ($\mu = 0.9 \times 10^{-8}$ per site per generation) (16), the estimated divergence at synonymous sites implies that the incompatible haplotypes diverged roughly 8 million generations ago.

Identification of *zeel-1*. We cloned *zeel-1* by introducing two fosmids carrying Bristol genomic DNA that together covered 45 kb of the 62-kb *zeel-1* interval into the Hawaii background (Fig. 2A). To test for rescue, we crossed transgenic individuals to Bristol and scored lethality among embryos from three crosses: self-fertilizing F_1 hermaphrodites, F_1 males backcrossed to Hawaii hermaphrodites, and Bristol × Hawaii F_1 males backcrossed to transgenic hermaphrodites. In the second cross, the transgene is inherited through the sperm, whereas in the third it is inherited through the oocyte. All crosses showed a reduction in embryonic lethality for transformants carrying fosmid B but not for those carrying fosmid A (Fig. 5). The ability of the transgene to mediate rescue when inherited through either sperm or oocyte implied that rescue occurs through zygotic transcription of *zeel-1*$_{Bristol}$. We genotyped surviving self-cross and male backcross progeny and found that survival of *zeel-1*$_{Hawaii}$ homozygotes required inheritance of the transgene, further supporting the idea that *zeel-1* acts zygotically (table S3).

To identify *zeel-1*, we individually subcloned the four predicted genes carried by fosmid B (Fig. 2A), introduced each into Hawaii, and tested for rescue. Only one subclone, containing predicted open reading frame *Y39G10AR.5*, rescued lethality, indicating that this gene is *zeel-1* (Fig. 5). *zeel-1* belongs to a previously uncharacterized *Caenorhabditis*-specific family of genes with homology to *zyg-11*, the substrate-recognition subunit of a CUL-2–based E3 ubiquitin ligase complex (17). *zeel-1* is located within the divergent interval and is deleted in Hawaii and all Hawaii-compatible wild isolates (table S2). To test whether

transgenic rescue requires ZEEL-1 protein, we generated a second *zeel-1* subclone identical to the first except that it contained a frame shift [via a 4–base pair (bp) insertion] truncating the protein at 25% of its length. When bombarded into Hawaii, the frameshifted transgene failed to rescue, indicating that rescue requires ZEEL-1 protein (Fig. 5).

Analysis of *peel-1*. Transgenic worms carrying Bristol-library fosmids, which together cover seven of the nine predicted genes in the *peel-1* interval, failed to induce the paternal-effect lethality of *zeel-1*$_{Hawaii}$ homozygotes, as did worms carrying subclones of the four genes within the divergent interval (8). Putatively null alleles of *srbc-64*, *nekl-1*, and *Y39G10AR.17*, as well as RNA interference (RNAi) targeting the remaining genes, did not abolish the paternal effect (tables S4 and S5). These negative results are equivocal because of potential germline silencing of transgenes, a possible requirement for chromosomal heterozygosity of *peel-1* in the spermatogenic germline, and the ineffectiveness of RNAi against sperm-expressed genes (18).

To find potential *peel-1* mutations, we examined the doubly compatible wild strain MY19 from Roxel, Germany. MY19 shows no lethality in crosses with Bristol and carries an intact *zeel-1* sequence, but it also fails to induce paternal-effect lethality in crosses with Hawaii, suggesting that it carries a Hawaii-like allele of *peel-1* or a suppressor of the paternal effect. To test for the existence of an unlinked suppressor, we mapped the inability of MY19 to induce lethality of *zeel-1*$_{Hawaii}$ homozygotes relative to a marker 10 cM from the *peel-1* interval (8). Absence of the paternal effect in MY19 mapped 10 cM from the marker (table S6), suggesting that MY19 might carry a mutation in *peel-1*. We sequenced all coding regions contained within the *peel-1* interval in MY19, as well as most noncoding regions located within the interval of elevated divergence between Bristol and Hawaii.

The MY19 haplotype is nearly identical to Bristol and contains only 12 nonsynonymous

polymorphisms, 4 of which fall in genes for which putative null alleles failed to abolish the paternal effect. Of the remaining polymorphisms, seven fall in *Y39G10AR.15*, a gene with spermatogenic germline expression (19) but located outside the divergent interval. The sequence of *Y39G10AR.16*, another gene with spermatogenic expression (19) but located within the divergent interval, contains no nonsynonymous polymorphisms. Notably, the coding sequence of *zeel-1*$_{MY19}$ is identical to that of *zeel-1*$_{Bristol}$, suggesting that the paternal effect does not arise from the activity of *zeel-1* itself or that if it does, MY19 and Bristol differ in their regulatory regions.

Discussion. We discovered a genetic incompatibility in *C. elegans* that causes lethality of embryos homozygous for a naturally occurring deletion of *zeel-1* when sired by individuals heterozygous for the Bristol and Hawaii alleles of a tightly linked paternal-effect gene, *peel-1*.

Paternal-effect mutations are rare, and only one has been described in *C. elegans* (20). The interaction between a paternal effect and a zygotically acting gene is surprising given that sperm-supplied factors are expected to act before zygotic transcription begins (9, 10). Zygotically expressed ZEEL-1 may be a molecular antidote to a sperm-carried PEEL-1 protein that is otherwise toxic during development. An analogous maternal effect by zygotic interaction has been described in *Tribolium* (21).

The Bristol haplotype of the *zeel-1/peel-1* interval gains a transmission advantage by inducing the lethality of embryos not inheriting it. This segregation distortion is characteristic of genic drive, in which selection at the level of alleles within an individual (genic selection) acts independently of selection at the level of individuals within a population (genotypic selection). Aside from this similarity, however, the *C. elegans* incompatibility does not conform to the expectations of genic drive. In drive systems, driver haplotypes are expected to fix within populations and are observed only where

Table 2. Genes in the interval of elevated divergence between the Bristol and Hawaii haplotypes exhibit signatures of purifying selection. d_N, nonsynonymous substitutions per nonsynonymous site. d_S, synonymous substitutions per synonymous site. ω, the d_N/d_S ratio. These quantities were estimated by maximum likelihood with the PAML program (35).

Gene	No. of sites	d_N	d_S	ω
		To the left of the divergent interval		
Y39G10AR.7	1422	0.001	0.000	–
mcm-4	2469	0.000	0.004	–
Y39G10AR.15	2112	0.001	0.000	–
		Inside the divergent interval		
srbc-64	870	0.018	0.147	0.124
Y39G10AR.16*	681	0.019	0.106	0.181
ugt-31	1563	0.022	0.158	0.140
		On the right boundary of the divergent interval		
nekl-1†	2949	0.001	0.019	0.050

*The Hawaii and Bristol sequences have different predicted splice sites on both sides of intron 3, resulting in several predicted amino acid residues that are not shared between the two alleles. We considered only the shared exonic sequences.   †nekl-1 cDNAs from Bristol and Hawaii differ slightly from the predicted exon structure. We used the exon structure from our cDNA clones (8).

Fig. 5. Transgenic complementation identifies *Y39G10AR.5* as *zeel-1*. Embryonic lethality was scored in three crosses: selfing F$_1$ hermaphrodites carrying the transgene, Hawaii hermaphrodites × F$_1$ males carrying the transgene, and Hawaii hermaphrodites carrying the transgene × F$_1$ males. Orange and blue bars designate Bristol and Hawaii haplotypes, respectively. Diagonal segments represent transgenes. Within crosses, each circle represents the result for an independent transgenic line (n/d, not determined). For each result, we scored at least 50 embryos; typically, 200 to 300 were scored. Solid circles mark results showing a significant reduction in lethality (one-sided χ^2, $P < 0.005$). Most transgenes were not integrated and were therefore not transmitted to all progeny. Arrows designate the single integrated line. In the male backcross with maternal transgene inheritance, where the integrated transgene was transmitted to all progeny, embryonic lethality was 4% ($n = 298$ embryos).

genotypic selection against them prevents fixation (*22*). Where this countervailing force is absent, drivers become fixed in a population and are detectable only in interpopulation crosses (*23–27*). In *C. elegans*, homozygotes of both haplotypes are fit, and incompatible haplotypes co-occur globally. One explanation for the long-term maintenance of both haplotypes is that genotypic selection favoring the Hawaii haplotype counterbalances genic drive favoring Bristol. Because *C. elegans* is primarily selfing and drive can influence transmission only in heterozygotes, this model would require a precise and stable level of outcrossing. Although a lack of knowledge of *C. elegans* population biology prevents us from rejecting this model, we favor an alternative explanation not subject to such constraints.

Whereas the drive model implies that balancing selection is a consequence of the incompatibility, an alternative is that the incompatibility is an incidental side effect of balancing selection. For example, balancing selection on pathogen resistance genes with costs in the absence of infection may maintain two haplotypes within a population for much longer than would be possible under neutral drift (*28*, *29*). If the haplotypes occur largely as homozygotes within a primarily selfing species, mutations arising on one haplotype are rarely tested against mutations on the other. Alleles fixed by drift or positive selection within the genetic context of their own haplotype but incompatible with alleles of another are not purged from the population but instead maintained alongside the balanced polymorphism,

which acts as an incompatibility trap. Because incompatible alleles are deleterious only after outcrossing, selection favoring the balanced polymorphism may be of lesser magnitude than selection against the incompatibility. Because long-term maintenance of incompatible alleles depends on tight linkage to the balanced polymorphism, the interacting loci are necessarily tightly linked to one another as well.

Our model, like genic drive, explains the occurrence of long-lived incompatible alleles at tightly linked loci, but it better incorporates the low level of outcrossing and the global co-occurrence of both haplotypes. Although the cause of balancing selection remains unknown, all genes within the divergent interval carry multiple nonsynonymous differences between the haplotypes, and several are known or predicted to interact with signals from the environment (*30–32*). The deletion of *zeel-1* may be analogous to presence/absence polymorphisms of pathogen resistance genes in *Arabidopsis*, which are known targets of balancing selection (*28*, *29*).

The *C. elegans* incompatibility suggests that long-term balancing selection in selfing species may facilitate the sympatric accumulation and maintenance of Dobzhansky-Muller type incompatibilities involving tightly linked loci. In the Dobzhansky-Muller model of speciation, incompatibilities emerge from the deleterious interactions of alleles that are neutral or advantageous in their own genetic backgrounds. Although classic models predict incompatible alleles to occur in allopatric populations (*33*), the *C. elegans* incom-

patibility occurs within interbreeding populations and does not appear to precipitate speciation, because gene flow between the incompatible classes occurs throughout the rest of the genome. The *C. elegans* incompatibility may be an example of incidental linkage between developmentally and ecologically important genes driving the evolution of development.

References and Notes
1. A. Barrière, M. A. Félix, *Curr. Biol.* **15**, 1176 (2005).
2. J. Hodgkin, T. Doniach, *Genetics* **146**, 149 (1997).
3. A. Barrière, M. A. Félix, *Genetics* **176**, 999 (2007).
4. A. Sivasundar, J. Hey, *Curr. Biol.* **15**, 1598 (2005).
5. D. R. Denver, K. Morris, W. K. Thomas, *Mol. Biol. Evol.* **20**, 393 (2003).
6. K. A. Swan *et al.*, *Genome Res.* **12**, 1100 (2002).
7. S. R. Wicks, R. T. Yeh, W. R. Gish, R. H. Waterston, R. H. Plasterk, *Nat. Genet.* **28**, 160 (2001).
8. See supporting material on *Science* Online.
9. S. W. L'Hernault, "Spermatogenesis" (20 February 2006), at WormBook, The *C. elegans* Research Community, doi/10.1895/wormbook.1.85.1 (www.wormbook.org).
10. P. Gonczy, L. S. Rose, "Asymmetric cell division and axis formation in the embryo" (15 October 2005), at WormBook, The *C. elegans* Research Community, doi/10.1895/wormbook.1.30.1 (www.wormbook.org).
11. M. Haber *et al.*, *Mol. Biol. Evol.* **22**, 160 (2005).
12. R. Koch, H. G. van Luenen, M. van der Horst, K. L. Thijssen, R. H. Plasterk, *Genome Res.* **10**, 1690 (2000).
13. A. Sivasundar, J. Hey, *Genetics* **163**, 147 (2003).
14. D. Charlesworth, *PLoS Genet.* **2**, e64 (2006).
15. A. D. Cutter, *Genetics* **172**, 171 (2006).
16. D. R. Denver, K. Morris, M. Lynch, W. K. Thomas, *Nature* **430**, 679 (2004).
17. S. Vasudevan, N. G. Starostina, E. T. Kipreos, *EMBO Rep.* **8**, 279 (2007).
18. J. Ahringer, "Reverse genetics" (6 April 2006), at WormBook, The *C. elegans* Research Community, doi/10.1895/wormbook.1.47.1 (www.wormbook.org).

19. V. Reinke, I. S. Gil, S. Ward, K. Kazmer, *Development* **131**, 311 (2004).
20. H. Browning, S. Strome, *Development* **122**, 391 (1996).
21. R. W. Beeman, K. S. Friesen, R. E. Denell, *Science* **256**, 89 (1992).
22. G. D. Hurst, J. H. Werren, *Nat. Rev. Genet.* **2**, 597 (2001).
23. A. Atlan, H. Mercot, C. Landre, C. Montchamp-Moreau, *Evol. Int. J. Org. Evol.* **51**, 1886 (1997).
24. R. W. Beeman, K. S. Friesen, *Heredity* **82**, 529 (1999).
25. L. Fishman, J. H. Willis, *Genetics* **169**, 347 (2005).
26. H. A. Orr, S. Irving, *Genetics* **169**, 671 (2005).
27. F. A. Reed, R. G. Reeves, C. F. Aquadro, *Evol. Int. J. Org. Evol.* **59**, 1280 (2005).
28. E. A. Stahl, G. Dwyer, R. Mauricio, M. Kreitman, J. Bergelson, *Nature* **400**, 667 (1999).
29. D. Tian, H. Araki, E. Stahl, J. Bergelson, M. Kreitman, *Proc. Natl. Acad. Sci. U.S.A.* **99**, 11525 (2002).
30. M. Shapira *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **103**, 14086 (2006).
31. J. H. Thomas, *Genome Res.* **16**, 1017 (2006).
32. J. H. Thomas, *Genetics* **172**, 127 (2006).
33. H. A. Orr, D. C. Presgraves, *Bioessays* **22**, 1085 (2000).
34. K. A. Frazer, L. Pachter, A. Poliakov, E. M. Rubin, I. Dubchak, *Nucleic Acids Res.* **32**, W273 (2004).
35. Z. Yang, *Comput. Appl. Biosci.* **13**, 555 (1997).
36. We thank the *Caenorhabditis* Genetics Center, the National Bioresource Project of Japan, the NemaGENETAG Consortium, M.-A. Félix, A. Barrière, E. Dolgin, and H. Van Epps for strains; R. Maruyama and A. Singson for advice; S. Skrovanek for lab assistance; and H. Coller, A. Cutter, D. Gresham, R. Gosh, L. Moyle, J. Shapiro, and E. Smith for comments on the manuscript. Supported by a National Defense Science and

REPORTS

Single-Molecule Cut-and-Paste Surface Assembly

S. K. Kufer,[1] E. M. Puchner,[1] H. Gumpp,[1] T. Liedl,[2] H. E. Gaub[1]

We introduce a method for the bottom-up assembly of biomolecular structures that combines the precision of the atomic force microscope (AFM) with the selectivity of DNA hybridization. Functional units coupled to DNA oligomers were picked up from a depot area by means of a complementary DNA strand bound to an AFM tip. These units were transferred to and deposited on a target area to create basic geometrical structures, assembled from units with different functions. Each of these cut-and-paste events was characterized by single-molecule force spectroscopy and single-molecule fluorescence microscopy. Transport and deposition of more than 5000 units were achieved, with less than 10% loss in transfer efficiency.

Functional biomolecular assembly aims to create structures from a large variety of biomolecular building blocks in a geometrically well-defined manner in order to create new functions (*1*, *2*), such as artificial signaling cascades or synergetic combinations of enzymes. Hybrid devices could include quantum dots co-assembled with dye molecules, or gold particles assembled as plasmon hot spots with a sample protein positioned into the focus (*3*). One way to assemble such molecular devices would be to physically pick up the different units needed with a scanning probe tip, translocate these units to a different location, and deposit them with high spatial precision (*4*–*6*). The entire process would also have to be carried out in an aqueous environment.

For the translocation of nanoscale objects, we used atomic force microscopy, which has been used in this context for mechanical single-molecule experiments (*7*–*12*) or lithography (*13*, *14*); however, previously suggested devices include the use of molecular pliers at the end of atomic force microscope (AFM) cantilevers that could grab and release the building blocks, triggered by an

external signal of either electrical or optical nature (*15*). We report a simpler and robust solution based on DNA hybridization and hierarchical bonds defined by different unbinding forces.

A well-sorted "depot," with a large variety of molecular species, stably stored in well-defined loci, is a prerequisite for the assembly of a multicomponent device. DNA chips offer a freely programmable pattern of oligomers that are commercially available and have spot sizes in the submicrometer range (*16*). Niemeyer *et al.* (*17*) converted such a DNA pattern into a protein pattern by binding a DNA-labeled protein to its corresponding spot on a DNA chip. The length of the oligomers can be chosen so that after incubation and stringent washing, a thermodynamically stable pattern of proteins is obtained. Given the known sequence map of the DNA chip, different molecular species can be stored in a known position on the depot chip. Alternatively, when only a limited variety of building blocks is needed, microfluidic elastomer channels may be used to create patterns (*18*–*20*) of building blocks, which after removal of the elastomer may be manipulated with the AFM tip (fig. S3).

We used this approach to store our functional units and also extended the DNA oligomers to fulfill a second function; namely, to serve as a handle (Fig. 1). This additional stretch of DNA

can hybridize to a complementary DNA covalently attached to an AFM tip. We chose the duplexes to be comparable in length and binding free energy, but we selected the sequences so that the anchor hybridizes in the so-called "unzip" geometry and the handle hybridizes in the "shear" geometry [Fig. 1 and (*21*)]. These two duplex geometries differ substantially in that, upon forced unbinding, the zipper duplex is opened up base pair by base pair, whereas in the shear geometry, all base pairs are loaded in parallel (Fig. 2 and fig. S1). Although the thermodynamic stability and the spontaneous off rate of both geometries are comparable, their rupture forces differ dramatically (*22*), as has been shown experimentally and was validated theoretically in several studies (*21*, *23*–*27*). Thus, upon retraction of the AFM tip, the anchor duplex will break open and the functional unit will be bound to the tip.

As can be seen in Fig. 2C, these force distance curves provide a characteristic fingerprint and serve as a robust criterion to decide whether a molecule was picked up from the depot. To avoid multiple transfers, we chose the density of the anchors on the tip to be low enough that in 35% of the attempts, only one unit was picked up, and in 20% of the attempts, just two units. In 20% of all attempts, we recorded traces like the lower two in Fig. 2C, which showed that we had not picked up any unit (fig. S5D). Because we recorded such a force distance curve for every pickup, we knew exactly how many units were transferred to the tip. The pickup process can be corrected online by either picking up more units or by dropping excess units in a "trash can" on the target area.

Once a unit is transferred to the tip, it can be moved to its new position on the target area. The target area had surface chemistry similar to that of the depot area, but the anchor oligomers were chosen so that when the tip was lowered, they bound to the transfer DNA in shear geometry and formed a duplex, which was longer than the handle duplex. Although the AFM tip can be positioned with subnanometer reproducibility, the precision with which the units can be

[1]Center for Nanoscience and Department of Physics, University of Munich, Amalienstrasse 54, 80799 Munich, Germany. [2]Department of Biological Chemistry and Molecular Pharmacology, Harvard Medical School, and Department of Cancer Biology, Dana-Farber Cancer Institute, Boston, MA 02115, USA.

deposited is limited by the lateral density of the anchor oligomers and their spacer length (presently in the 10-nm range).

Upon retraction of the tip from the surface, the force in the two DNA duplexes in series gradually increases until the weaker of the two complexes ruptures. The upper traces in Fig. 2D show examples of this process, which differs considerably in its signature from the unzipping shown in Fig. 2C (28, 29). It was shown (21) that a length difference of 10 base pairs (bp) is sufficient to make the rupture of the shorter handle duplex more likely by one order of magnitude than the rupture of the longer anchor duplex. As was the case during pickup, no bond rupture was detected in certain cases (Fig. 2D, lower traces). Here the hybridization with the target anchor oligomer had failed, although a functional unit was offered. In the majority of cases, a second or third attempt made a few nanometers away from the originally planned target spot was successful. Again, a protocol with a characteristic force distance curve (for brevity referred to as a transfer protocol) was recorded for each transfer event. After delivery of the functional unit to the target, the oligomer co-valently attached to the AFM tip was free again to hybridize with another handle sequence in the depot area.

For simplicity, we transferred only functional units of the same species but created patterns from single units with multiple functions instead. As functional units, we used molecular constructs consisting of a fluorophore (rhodamine green), a generic small ligand (biotin), and a DNA with extra overlap (which allows further docking of other units to the assembly in a later step) (Fig. 2 shows the schematics). The units were picked up
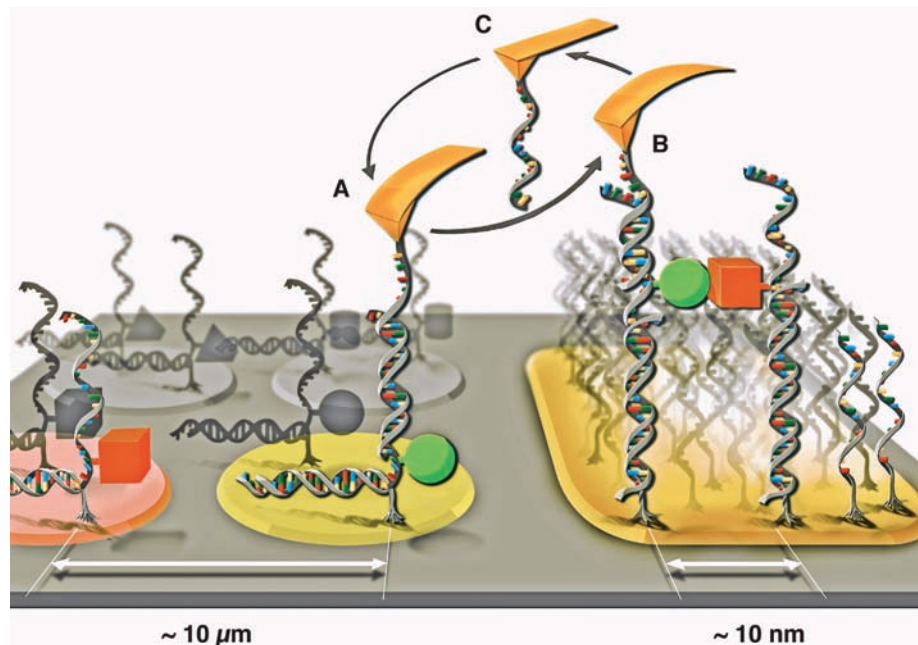
Fig. 1. Cartoon of the single-molecule cut-and-paste process. (**A**) Individual nanosized objects are picked up from discrete storage sites with a DNA oligomer at the tip of an AFM cantilever and transferred to a target site, where they are deposited with high spatial precision (**B**). (**C**) The length and binding geometry of the oligomers, which are used as an anchor or a handle, are chosen so that a hierarchy of unbinding forces allows the repetition of this process over and over again.

Fig. 2. Design of the assembled pattern with typical transfer protocols. (**A**) Individual functional units stored on the depot were picked up one at a time and transferred to the target area. The functional units consisted of a DNA oligomer with anchor and handle sequences, one fluorophore, one biotin, and an additional DNA binding site. (**B**) Five different patterns with different geometries were assembled on spots, which were 2 μm apart. In the first spot, we deposited one unit; in the second spot, two units with a lateral spacing of 50 nm; in the third spot, three units, and so forth. The lateral precision of the closed-loop feedback was ±6 nm. Force distance curves were recorded in every cycle as transfer protocols recording successful pickup and deposition. (**C**) Typical force distance curves measured during the pickup of functional units from the depot. At low extensions, the cantilever acts against the entropic force of the polyethylene glycol—DNA complex. When the force reaches about 20 pN, the anchor sequence is pulled open in a zipperlike mode, resulting in a plateau ~20 nm long. (**D**) Typical force distance curves recorded during the deposition of a single unit to the target area. The shape, with its sudden drop in force at about 50 pN, is characteristic of a rupture of a 20-bp DNA handle duplex loaded in shear geometry. The lower two force distance curves in (C) and (D) show attempts, where no transfer had occurred.
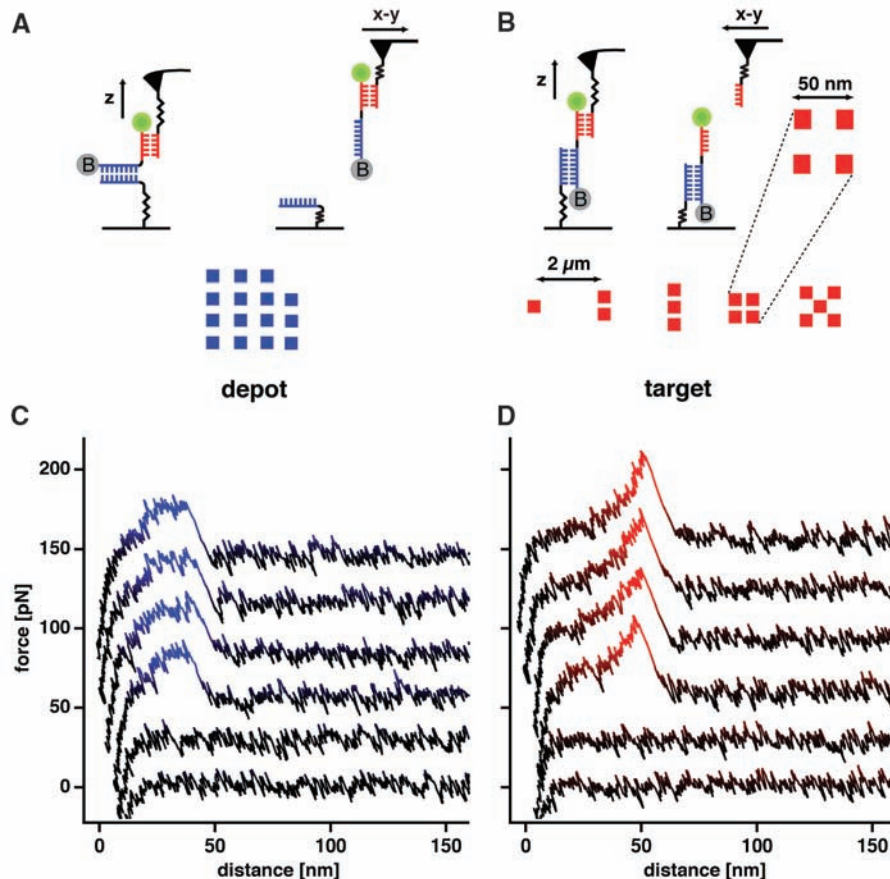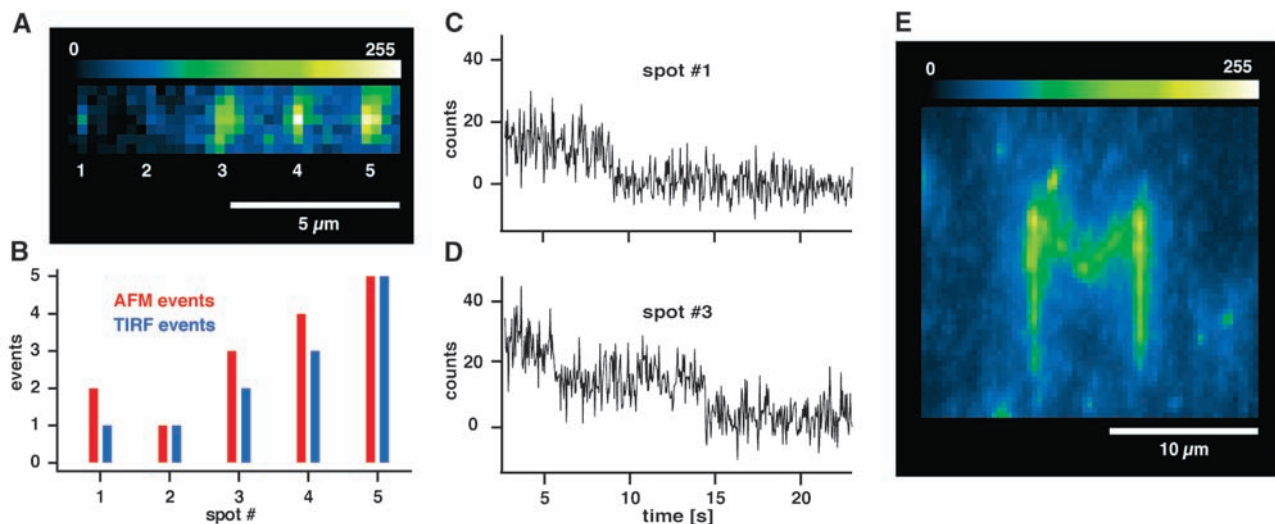
Fig. 3. (A) Fluorescence micrograph of the assembled pattern described in Fig. 2, imaged with TIRF excitation. The image was averaged over 440 frames, with 50 ms of integration time each. Because of the diffraction limit, individual fluorophores cannot be resolved spatially but can be resolved temporally. **(C)** and **(D)** show time traces of the two diffraction-limited 3 × 3–pixel–sized spots 1 and 3, exhibiting the typical stepwise bleaching of one and two single fluorophores, respectively. **(B)** Correlation analysis of the number of successfully transferred units as judged by the AFM transfer protocols and the number of bleaching steps in the fluorescence. **(E)** The capital letter M written by transporting 400 molecules from the depot area to defined positions at the target area. It was assembled with a tip that had already been used to transport more than 5000 molecules from the depot to the target area. These results demonstrate the long-term stability of the tip functionalization and the possibility of assembling extended constructs.

from a 100-nm-square pattern (Fig. 2A) and transferred to the target area. A pattern, as sketched in Fig. 2B, was assembled where we deposited a single unit in the first spot, a pair in the second spot, a triplet in the third, and so on. The transfer protocols always documented the actual number of transferred units.

After the assembly was completed, total internal reflection fluorescence (TIRF) imaging (30) showed discrete spots at the predicted positions (Fig. 3A). Because of the limited optical resolution, no details of the assemblies are resolved, but larger assemblies appear brighter. Time traces (Fig. 3, C and D) exhibited well-pronounced steps that were a clear indication of bleaching of individual fluorophores (31, 32) (movie S1). On the spot in the first column, we recorded a single step only, and the fluorophore was bleached after 9 s. The spot in the third column exhibited only two steps, although our transfer protocol reported the deposition of three functional units. Either one of the fluorophores was inactive from the beginning or it was bleached during the first 2 s of the illumination and not recorded because of background fluorescence. A direct correlation between the number of deposited units as judged by the transfer protocol and the number of bleaching steps is given in Fig. 3B. Both independent experiments are in excellent agreement, which indicates that we lost only a minor fraction of fluorophores in the transfer process.

In order to demonstrate the formation of larger constructs, we assembled the capital letter M shown in Fig. 3E. It consists of 400 units and was written with a tip that had already been used to transport more than 5000 functional units from the depot to the target area. All of the data shown here and in the supporting online material were recorded with one cantilever. Because the pickup probability dropped by only 10% toward the end of the experiment, the lifetime of the tip functionalization was adequate. The pattern was assembled at an average rate of 7 s per transfer. This slow transfer rate was limited by the rather large distance between depot and target area of 15 μm and the closed-loop feedback of the translational stages of the instrument. The online analysis of the transfer protocols was also not optimized. The physical limits are given by the resonance frequency of the piezo stage, and so improvements of several orders of magnitude are possible (33). With the development of massively parallel–operating AFM cantilevers (34), molecule-by-molecule assembly based on hierarchical forces may evolve into a versatile technology.

References and Notes

1. B. A. Grzybowski, H. A. Stone, G. M. Whitesides, *Nature* **405**, 1033 (2000).
2. T. Pellegrino *et al.*, *Small* **1**, 48 (2005).
3. M. Ringler *et al.*, *Nano Lett.* **7**, 2753 (2007).
4. D. M. Egler, E. K. Schweizer, *Nature* **344**, 524 (1990).
5. M. T. Cuberes, R. R. Schlittier, J. K. Gimzewski, *Appl. Phys. Lett.* **69**, 3016 (1996).
6. S. J. Greissl *et al.*, *J. Phys. Chem. B* **108**, 11556 (2004).
7. G. Binnig, H. Rohrer, C. Gerber, E. Weibel, *Phys. Rev. Lett.* **49**, 57 (1982).
8. G. Binnig, C. F. Quate, C. Gerber, *Phys. Rev. Lett.* **56**, 930 (1986).
9. P. K. Hansma, V. B. Elings, O. Marti, C. E. Bracker, *Science* **242**, 209 (1988).
10. M. Radmacher, R. W. Tillmann, M. Fritz, H. E. Gaub, *Science* **257**, 1900 (1992).
11. P. E. Marszalek *et al.*, *Nature* **402**, 100 (1999).
12. D. Fotiadis *et al.*, *Curr. Opin. Struct. Biol.* **16**, 252 (2006).
13. M. Jaschke *et al.*, *Biosens. Bioelectron.* **11**, 601 (1996).
14. R. D. Piner, J. Zhu, F. Xu, S. Hong, C. A. Mirkin, *Science* **283**, 661 (1999).
15. C. P. Collier *et al.*, *Science* **289**, 1172 (2000).
16. S. P. Fodor *et al.*, *Science* **251**, 767 (1991).
17. C. M. Niemeyer, T. Sano, C. L. Smith, C. R. Cantor, *Nucleic Acids Res.* **22**, 5530 (1994).
18. E. Delamarche, A. Bernard, H. Schmid, B. Michel, H. Biebuyck, *Science* **276**, 779 (1997).
19. S. R. Quake, A. Scherer, *Science* **290**, 1536 (2000).
20. D. C. Duffy, J. C. McDonald, O. J. A. Schueller, G. M. Whitesides, *Anal. Chem.* **70**, 4974 (1998).
21. C. Albrecht *et al.*, *Science* **301**, 367 (2003).
22. A hand-waving argument: Upon separation, the binding energy is overcome in the shear geometry within a much shorter distance than in the unzip geometry, therefore the force to overcome the energy barrier is much lower in the unzip geometry. Because the forced unbinding of the oligomer in shear geometry is a nonequilibrium process, its unbinding force is rate-dependent. In all experiments shown here, the duplexes were loaded with a rate of 3000 pN/s.
23. G. I. Bell, *Science* **200**, 618 (1978).
24. S. B. Smith, Y. J. Cui, C. Bustamante, *Science* **271**, 795 (1996).
25. B. Essevaz-Roulet, U. Bockelmann, F. Heslot, *Proc. Natl. Acad. Sci. U.S.A.* **94**, 11935 (1997).
26. M. Rief, H. Clausen-Schaumann, H. E. Gaub, *Nat. Struct. Biol.* **6**, 346 (1999).
27. G. Neuert, C. H. Albrecht, H. E. Gaub, *Biophys. J.* **93**, 1215 (2007).
28. J. Morfill *et al.*, *Biophys. J.* **93**, 2400 (2007).
29. T. Strunz, K. Oroszlan, R. Schafer, H. J. Guntherodt, *Proc. Natl. Acad. Sci. U.S.A.* **96**, 11277 (1999).
30. R. D. Vale *et al.*, *Nature* **380**, 451 (1996).
31. G. Seisenberger *et al.*, *Science* **294**, 1929 (2001).
32. P. Tinnefeld, M. Sauer, *Angew. Chem. Int. Ed.* **44**, 2642 (2005).
33. G. Schitter *et al.*, *IEEE Trans. Control Syst. Technol.* **15**, 906 (2007).
34. P. Vettiger *et al.*, *IBM J. Res. Devel.* **44**, 323 (2000).
35. Helpful discussions with P. Hansma, G. M. Whitesides, J. Fernandez, H. Heus, P. Tinnefeld, J. Morfill, C. Albrecht, and L. Whetton are gratefully acknowledged. Supported by the German Science Foundation and the Nanosystems Initiative Munich.

Electronic Liquid Crystal State in the High-Temperature Superconductor YBa$_2$Cu$_3$O$_{6.45}$

V. Hinkov,[1]* D. Haug,[1] B. Fauqué,[2] P. Bourges,[2] Y. Sidis,[2] A. Ivanov,[3] C. Bernhard,[4] C. T. Lin,[1] B. Keimer[1]

Electronic phases with symmetry properties matching those of conventional liquid crystals have recently been discovered in transport experiments on semiconductor heterostructures and metal oxides at millikelvin temperatures. We report the spontaneous onset of a one-dimensional, incommensurate modulation of the spin system in the high-transition-temperature superconductor YBa$_2$Cu$_3$O$_{6.45}$ upon cooling below ~150 kelvin, whereas static magnetic order is absent above 2 kelvin. The evolution of this modulation with temperature and doping parallels that of the in-plane anisotropy of the resistivity, indicating an electronic nematic phase that is stable over a wide temperature range. The results suggest that soft spin fluctuations are a microscopic route toward electronic liquid crystals and that nematic order can coexist with high-temperature superconductivity in underdoped cuprates.

The electronic states near the Fermi level of high-temperature superconductors derive from the hybridized d- and p-orbitals of copper and oxygen ions in a square-planar network. At a doping level of 1/8 hole per Cu ion, experimental work on a specific superconducting cuprate family, (La,Nd)$_{2-x}$(Sr,Ba)$_x$CuO$_4$ (La214), has shown that the two-dimensional (2D) electron system in the CuO$_2$ layers can support a state with uniaxial spin (*1–3*) and charge (*1, 4, 5*) order ("stripes"). Static stripe order implies that both translational and rotational symmetries of the copper-oxide square lattice are spontaneously broken. More unusual "electronic liquid crystal" states (*6*) that break the rotational symmetry of the lattice while at least partially preserving its translational symmetry can arise from quantum fluctuations of stripes (*6–8*), or from Fermi surface instabilities (*9–11*). Electronic nematic states have recently been discovered in semiconductor heterostructures (*12*) and in the bulk transition metal oxide Sr$_3$Ru$_2$O$_7$ (*13*). In both cases, however, they are stable only at millikelvin temperatures and in high magnetic fields and have thus far only been probed by transport measurements. We use neutron scattering to address the role of magnetic degrees of freedom in driving the formation of electronic liquid crystals and to explore the presence of liquid-crystalline order in the cuprates.

In the previous experiments on semiconductor heterostructures and on Sr$_3$Ru$_2$O$_7$, the nematic director was aligned by external magnetic fields, resulting in a strong macroscopic anisotropy of the current flow (*12, 13*), analogous to the alignment of nematic domains in conventional liquid crystals by electric fields or confining walls. In the cuprates, subtle crystallographic distortions can serve as aligning fields for symmetry-broken electronic phases. They reduce the four-fold rotational symmetry of the CuO$_2$ layer to a two-fold rotational or mirror symmetry by introducing a slight (~1%) difference between the in-plane lattice parameters. In stripe-ordered La214, the stripe domains in every CuO$_2$ layer are aligned by such a two-fold axis, but the layers are stacked in such a way that the global crystal symmetry is tetragonal and no macroscopic anisotropy is observed (*1*). The situation is more favorable in metallic YBa$_2$Cu$_3$O$_{6+x}$ (Y123), where a macroscopic orthorhombic crystal structure is stabilized by 1D CuO chains lying between the CuO$_2$ bilayers (*14*). Indeed, an unexpectedly large temperature- and doping-dependent in-plane anisotropy of the resistivity has been reported for this material (*15*). Although this constitutes possible evidence of nematic order (*7, 15*), clear signatures of an isotropic-to-nematic transition are not apparent in the transport data.

A determination of the spin-spin correlation function by magnetic neutron scattering has the potential to build a more robust case for nematic order and to elucidate the underlying microscopic mechanisms. Our sample is a large array of untwined YBa$_2$Cu$_3$O$_{6.45}$ single crystals with superconducting transition temperature T_c = 35 K (*14, 16*). In this work, we show only magnetic excitations that are odd under the exchange of the two layers within a CuO$_2$ bilayer (*14, 17*). Figure 1, A to C, provides an overview of the magnetic spectrum along the two perpendicular in-plane axes *a** and *b** at a temperature of 5 K. As previously observed for Y123 samples at similar doping levels (*17–20*), the magnetic intensity is concentrated around the in-plane wave vector $\mathbf{Q}_{AF} = (\pi/a, \pi/b)$ [the propagation vector of the antiferromagnetic state in undoped Y123, corresponding to (0.5, 0.5) in reciprocal lattice units (r.l.u.) (*14*)] and the spectrum is nearly gapless (Fig. 3C). The new aspect uncovered by the experiments on twin-free samples presented here is the in-plane anisotropy of the spectrum, which exhibits an unusual evolution with energy and temperature. Whereas at high excitation energies, *E*, the spectrum is isotropic (Fig. 1C), maps of the magnetic spectral weight for *E* < 15 meV reveal a pronounced anisotropy that increases with decreasing energy (Fig. 1, A and B). This anisotropy could not be recognized in previous work because of crystal twinning (Fig. 1D). Cuts through the spectrum with high instrumental resolution (Fig. 1, E and F) reveal that the anisotropic intensity distribution is generated by two incommensurate peaks symmetrically displaced from $\mathbf{Q}_{AF}$ along *a**, whereas along *b** the distribution is commensurate.

Figure 2 shows the temperature dependence of constant-energy cuts along the two high-symmetry axes. Although the isotropic intensity distribution at high *E* is temperature-independent, the anisotropy at low *E* is strongly reduced with increasing temperature. Fits of the profiles to two Gaussians centered at $\mathbf{Q} = \mathbf{Q}_{AF} \pm (\delta, 0)$ yield excellent descriptions of the data at all temperatures (lines in Fig. 1E and Fig. 2). The results of the fits for E = 3 meV are plotted in Fig. 3. The intrinsic half-widths-at-half-maximum along *a** and *b**, ξ_a^{-1} and ξ_b^{-1} (extracted from the Gaussian peaks after a correction for the instrumental resolution) are nearly identical and weakly temperature dependent. The zero-temperature offset suggests that magnetic long-range order is absent, reflecting the influence of magnetic quantum fluctuations, low-energy charge fluctuations, and/or disorder.

The incommensurability δ exhibits an order-parameter–like behavior with an onset temperature ~150 K (Fig. 3A). Although δ can no longer be accurately determined (nor is it physically meaningful) once it becomes smaller than ξ_a^{-1} (shaded area in Fig. 3A), this behavior indicates an underlying phase transition where $\delta \to 0$. At the same temperature, the $\mathbf{Q}$-integrated spin susceptibility $\chi''(\omega)$ exhibits a strong upturn (Fig. 3B), which is also reminiscent of a broadened phase transition. A rounding of the singularities related to a phase transition is generally expected because the spin system is probed at nonzero energy, but additional contributions can arise from disorder and/or the orthorhombic distortion of the crystal structure.

Ordinary magnetic phase transitions are associated with the formation of static magnetic moments. We have performed zero-field muon-spin-relaxation (μSR) experiments capable of routinely detecting static or slowly fluctuating electronic moments on the order of 0.01 μ$_B$ per lattice site. Figure 4A shows that at temperatures above 10 K, the relaxation of muon spins implanted into our samples is entirely determined

[1]Max-Planck-Institut für Festkörperforschung, Heisenbergstrasse 1, D-70569 Stuttgart, Germany. [2]Laboratoire Léon Brillouin, Commissariat à L'Énergie Atomique, Centre National de la Recherche Scientifique, CEA-Saclay, F-91191 Gif-sur-Yvette, France. [3]Institut Laue-Langevin, 6 Rue Jules Horowitz, F-38042 Grenoble cedex 9, France. [4]Department of Physics and FriMat Center for Nanomaterials, University of Fribourg, Chemin du Musée 3, CH-1700 Fribourg, Switzerland.

*To whom correspondence should be addressed. E-mail: v.hinkov@fkf.mpg.de
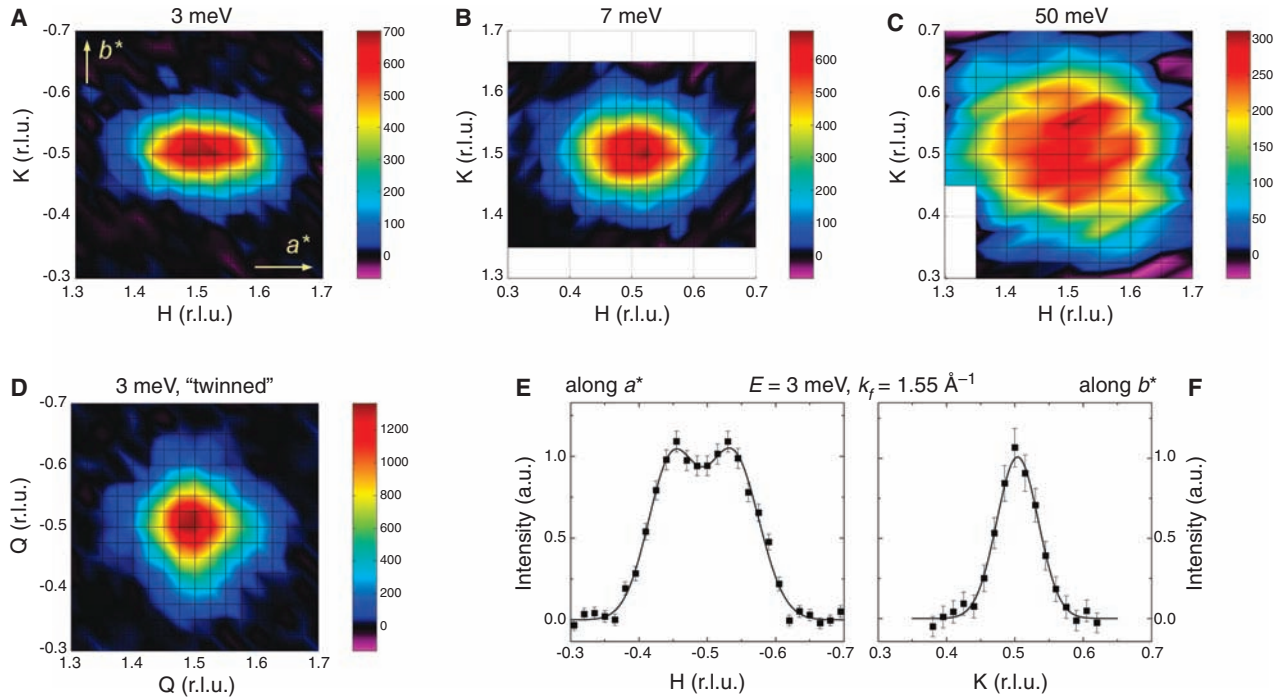
Fig. 1. Geometry of the spin excitations around Q_{AF} at L = 1.7 r.l.u. in the a-b plane at $T = 5$ K. (**A** to **C**) Intensity maps of the spin-excitation spectrum at 3, 7, and 50 meV, respectively, assembled from triple-axis scans. The a^* and b^* directions are indicated in (A). The scale is in arbitrary units and the wave vector is in r.l.u. (14). The wave vector of the scattered neutrons, k_f, was fixed to 2.66 Å^{-1} in (A) and (B) and to 4.1 Å^{-1} in (C). The crossings of black lines represent measured data points. All scans were corrected for a **Q**-linear background. (**D**) Color map of the intensity at 3 meV, as it would be observed in a crystal consisting of two perpendicular twin domains with equal population. The representation was obtained by transposing the map in (A) and superposing it with the original map. (**E** and **F**) Scans along a^* and b^* through Q_{AF}. The resolution was enhanced as compared with (A) to (D) by working with $k_f = 1.55$ Å^{-1}. Solid lines represent fits with one (F) or two (E) Gaussians to the data. A linear background was subtracted.

Fig. 2. Temperature evolution of the a-b anisotropy of the spin correlations. Full squares and empty circles represent data points measured at fixed K along a^* and at fixed H along b^*, respectively. (**A** to **C**) $E = 3$ meV; (**D** to **F**) $E = 7$ meV; (**G** to **I**) $E = 25$ meV. The measurements were performed at $T = 5$ K in (A), (D), and (G); at $T = 40$ K in (B), (E), and (H); and at $T = 100$ K in (C), (F), and (I). The final wave vector k_f was fixed to 2.66 Å^{-1}. All scans are normalized to unity to allow a better comparison of the scan profiles. Solid lines represent the results of fits with one or two Gaussians. A linear background was subtracted.
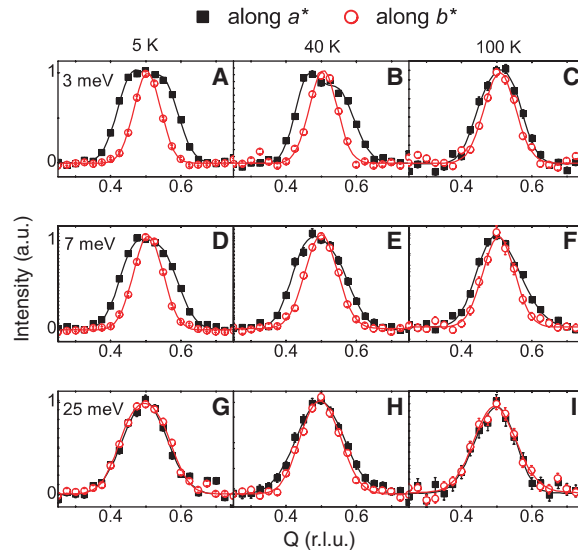
by nuclear spins. In agreement with previous work (21), an additional contribution to the μSR signal from low-energy electronic spin excitations is seen below 10 K. Manifestations of electronic magnetic moments that are static on the microsecond time scale probed by the muons are discernible only in the decay profiles below ~2 K.

Information about the spatial structure of these low-energy spin correlations was obtained by additional neutron-scattering measurements with a spectrometer nominally set for zero energy transfer. As a result of the finite energy resolution, spin fluctuations with energies below the instrumental resolution width of ~0.2 meV are probed in this "quasi-elastic" configuration. Representative data (Fig. 4B) show that the incommensurability of the quasi-elastic profile as well as its width ξ_a^{-1} are nearly identical to those of the low-energy inelastic data plotted in Fig. 3A; ξ_b^{-1} is somewhat smaller. The quasi-elastic intensity exhibits a significant upturn below 30 K. These data fit well into the "spin-freezing" phenomenology of deeply underdoped cuprates (20–24). The characteristic frequency of electronic spin fluctuations is reduced with decreasing temperature; this is directly apparent in the inelastic neutron-scattering data of Fig. 3C. Upon further cooling, the spin system gradually freezes into an ensemble of slowly fluctuating, finite-sized domains. The characteristic fluctuation rate of these domains progressively enters the frequency windows of quasi-elastic neutron scattering (20, 22, 23), μSR (21, 22), and nuclear magnetic resonance (24) methods.

The results demonstrate that the spin correlations within the fluctuating domains are incommensurate. Because the signal derives entirely from the CuO$_2$ planes (14), coupling between spins in the planes and the CuO chains can be ruled out as a substantial driving force of these correlations. Thus, our data demonstrate that the spin system in the CuO$_2$ planes of strongly underdoped Y123 becomes inherently unstable toward the formation of a uniaxial, slowly fluctuating spin texture at a critical temperature of ~150 K. The small (<1%) orthorhombic distortion of the crystal structure serves as an aligning field for the incommensurate domains and leads to the large anisotropy of the neutron-scattering pattern below 150 K. This scenario is also consistent with the observation of local uniaxial charge domains in the spin-glass state of other high-temperature superconductors by scanning tunneling spectroscopy (25).
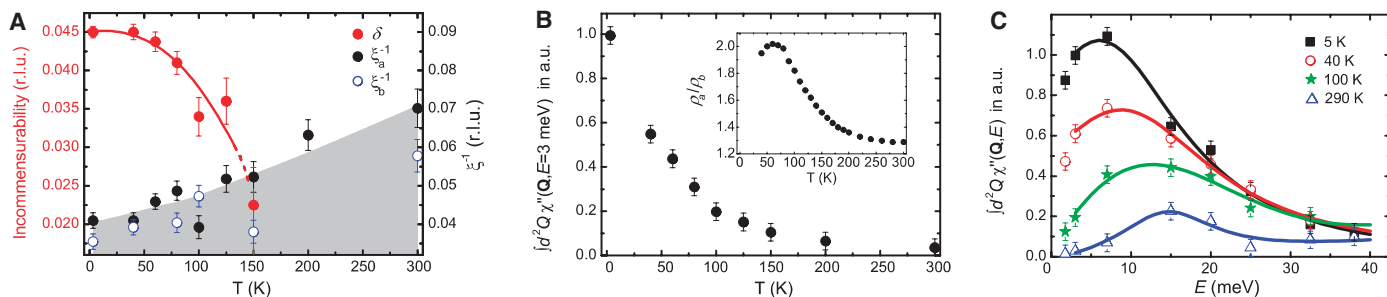
Fig. 3. Temperature and energy evolution of parameters characterizing the spin excitation spectrum. The parameters are the results of fits to the raw data, corrected for the instrumental resolution. (**A**) Incommensurability δ (red symbols), half-width-at-half-maximum of the incommensurate peaks along a^* (ξ_a^{-1}, black symbols) and along b^* (ξ_b^{-1}, open blue symbols) in reciprocal lattice units. The upper border of the shaded area follows ξ_a^{-1}. The ξ^{-1} axis is scaled to twice the value of the δ axis, hence δ-points lying inside the shaded area indicate an incommensurate peak separation below ξ_a^{-1}. (**B**) Imaginary part of the **Q**-integrated spin susceptibility $\chi''(\omega)$ at 3 meV. The inset shows the ratio of the electrical resistivity along a^* and b^* of a sample with similar doping levels as ours, reproduced from (*15*). (**C**) Energy evolution of $\chi''(\omega)$ at $T = 5, 40, 100,$ and 290 K. In all panels, the error bars were estimated from the fits. The lines are a guide to the eye.

Fig. 4. Measurements of the quasi-elastic magnetic response. (**A**) Zero-field muon-spin-relaxation data taken at $T = 1.5$, 2.75, 3.5, 5, 7, and 30 K. Solid lines show the results of fits to a relaxation function that consists of the product of a stretched exponential and a Kubo-Toyabe function, which account for the contribution of the electronic and nuclear magnetic moments, respectively. (**B** and **C**) Neutron-scattering scans along a^* at fixed $K = -0.5$ and along b^* at fixed $H = 0.54$, with a nominal energy transfer of 0 meV. The difference between the intensities at 5 K and 40 K is shown in arbitrary units, and k_f was fixed to 1.48 Å^{-1}. Solid lines represent fits with one (C) or two (B) Gaussians to the data. A linear background was subtracted.
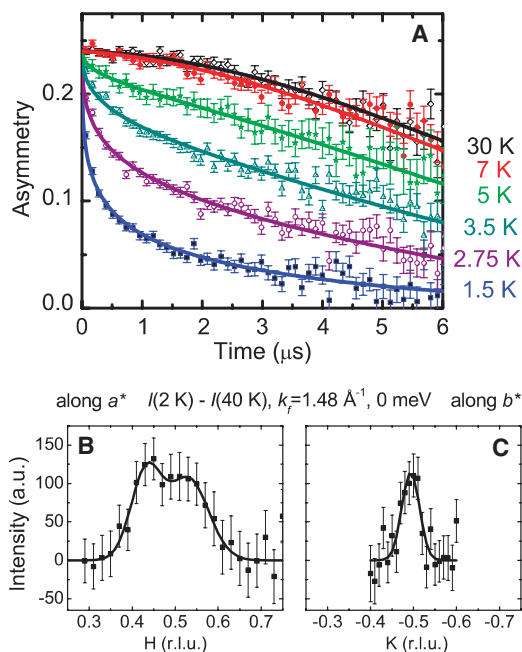
The relationship between the magnetic dynamics and the charge transport properties of YBa$_2$Cu$_3$O$_{6.45}$ can be discussed in the context of theoretical proposals for electronic nematic order. Although in-plane anisotropies of both the spin fluctuation spectrum and the electrical resistivity in the absence of spin and charge order are generic features of nematic states in theoretical models of quasi-2D electron systems (*10*, *11*), nematic order has thus far been diagnosed solely on the basis of the spontaneous onset of anisotropic resistivity (*12*, *13*). In Y123, the case for nematicity is harder to make based on transport alone, because the aligning field is not tunable like the magnetic fields used in (*12*, *13*). The resistivity is thus anisotropic at all temperatures, and the putative isotropic-to-nematic transition is inevitably broadened. The pronounced enhancement of the in-plane resistivity ratio, ρ_a/ρ_b, below ~200 K (inset in Fig. 3B) is, however, suggestive of an underlying phase transition (*15*). The strikingly similar temperature dependence of the spectral weight of the anisotropic low-energy spin fluctuations (Fig. 3B) confirms this interpretation. In addition, both quantities exhibit a parallel evolution with energy and doping: Optical conductivity measurements on Y123 crystals with doping levels similar to ours indicate that the charge transport anisotropy is strongly reduced at excitation energies above ~20 meV (*26*), in good agreement with the crossover to isotropic spin fluctuations we observe. In more highly doped Y123 (where the orthorhombic distortion is enhanced), ρ_a/ρ_b is reduced and much less temperature dependent (*15*), whereas the spectral weight of the low-energy collective spin excitations is strongly diminished (*17–19*, *27–29*). Although low-energy spin excitations are not entirely suppressed, this phenomenon has been termed a "spin-gap." We note that the dispersion and in-plane geometry of the excitations at energies exceeding the spin-gap of YBa$_2$Cu$_3$O$_{6+x}$ ($x \geq 0.5$) above T_c bear some resemblance to the low-energy excitations of YBa$_2$Cu$_3$O$_{6.45}$ (*19*, *28*, *29*). This suggests the presence of a quantum phase transition at $x \sim 0.5$, where nematic order disappears and a gap opens up in the spin fluctuation spectrum. Whereas the spin fluctuations above the gap remain characteristic of the ordered phase nearby, the dc-transport properties appear to be primarily controlled by interactions with low-energy excitations.

The energy and momentum dependence of the spin excitations we have observed helps to develop a microscopic description of the nematic state and to discriminate between different theoretical descriptions of the coupling between spin and charge excitations in the cuprates. Specifically, the data of Fig. 3A are incompatible with theories according to which the resistivity anisotropy is controlled by an anisotropic spin-spin correlation length (*30*). Rather, the data indicate a correspondence between the upturn in ρ_a/ρ_b and the onset of the incommensurate modulation of the spin system. Incommensurate peaks in the magnetic neutron-scattering pattern can arise either from a longitudinal modulation, in which the magnetic moments are collinear but their amplitude is spatially modulated, or from a transverse modulation, in which the moment direction varies but the amplitude remains constant. Our data are compatible with slow fluctuations characteristic of either type of modulation. Spin-amplitude modulated states naturally go along with a modulation of the charge carrier density, and the carrier mobilities along and perpendicular to the modulation axis are generally expected to be different (*7*). It has also been shown that a transverse modulation with spiral spin correlations can lead to anisotropic hopping transport in weakly doped cuprates with diverging low-temperature resistivity (*31*). Further work is required to assess whether this mechanism can be generalized to metallic electron systems such as the one in YBa$_2$Cu$_3$O$_{6.45}$. Finally, we note that the magnitude of δ in YBa$_2$Cu$_3$O$_{6.45}$ is incompatible (*16*, *18*) with the Yamada-plot (*32*), that is, the δ-versus-x relation that holds generally for

La214. This relation stipulates an incommensurability $\delta \sim 0.085$ r.l.u. for our doping level of 0.085 holes per Cu atom, whereas we observe δ to saturate at ~ 0.045 r.l.u. at the lowest energies and temperatures. Besides, a spontaneous onset of the incommensurability has thus far not been observed in this cuprate family (2).

The spin dynamics of $YBa_2Cu_3O_{6.45}$ and its close correspondence with the charge transport properties of this material provide strong evidence of a cooperative transition to an electronic nematic phase at a temperature that is about two orders of magnitude higher than the onset of static magnetic order. In the light of previous work on La214 (33), we expect the quasistatic, incommensurate magnetic order observed here to be enhanced by a magnetic field perpendicular to the CuO_2 planes, with possibly important consequences for the interpretation of the Fermi surface pockets inferred from recent high-field quantum oscillation measurements on Y123 crystals with doping levels similar to ours (34, 35).

References and Notes

1. J. M. Tranquada, B. J. Sternlieb, J. D. Axe, Y. Nakamura, S. Uchida, *Nature* **375**, 561 (1995).
2. M. Fujita, H. Goka, K. Yamada, J. M. Tranquada, L. P. Regnault, *Phys. Rev. B* **70**, 104517 (2004).
3. N. B. Christensen *et al.*, *Phys. Rev. Lett.* **98**, 197003 (2007).
4. P. Abbamonte *et al.*, *Nat. Phys.* **1**, 155 (2005).
5. M. V. Zimmermann *et al.*, *Europhys. Lett.* **41**, 629 (1998).
6. S. A. Kivelson, E. Fradkin, V. J. Emery, *Nature* **393**, 550 (1998).
7. S. A. Kivelson *et al.*, *Rev. Mod. Phys.* **75**, 1201 (2003).
8. V. Cvetkovic, J. Zaanen, *Phys. Rev. Lett.* **97**, 045701 (2006).
9. V. Oganesyan, S. A. Kivelson, E. Fradkin, *Phys. Rev. B* **64**, 195109 (2001).
10. Y.-J. Kao, H.-Y. Kee, *Phys. Rev. B* **72**, 024502 (2005).
11. H. Yamase, W. Metzner, *Phys. Rev. B* **73**, 214517 (2006).
12. K. B. Cooper, M. P. Lilly, J. P. Eisenstein, L. N. Pfeiffer, K. W. West, *Phys. Rev. B* **65**, 241313 (2002).
13. R. A. Borzi *et al.*, *Science* **315**, 214 (2007).
14. Materials and methods, as well as supporting text, are available as supporting material on *Science* Online.
15. Y. Ando, K. Segawa, S. Komiya, A. N. Lavrov, *Phys. Rev. Lett.* **88**, 137005 (2002).
16. The lattice parameters, $a = 3.8388$ Å, $b = 3.8747$ Å, and $c = 11.761$ Å, correspond to the chemical composition $YBa_2Cu_3O_{6.45}$ and a hole concentration of 0.085 ± 0.01 per copper ion (36).
17. H. F. Fong *et al.*, *Phys. Rev. B* **61**, 14773 (2000).
18. P. Dai, H. A. Mook, R. D. Hunt, F. Dogan, *Phys. Rev. B* **63**, 054525 (2001).
19. C. Stock *et al.*, *Phys. Rev. B* **69**, 014502 (2004).
20. C. Stock *et al.*, *Phys. Rev. B* **73**, 100504(R) (2006).
21. C. Niedermayer *et al.*, *Phys. Rev. Lett.* **80**, 3843 (1998).
22. B. J. Sternlieb *et al.*, *Phys. Rev. B* **41**, 8866 (1990).
23. B. Keimer *et al.*, *Phys. Rev. B* **46**, 14034 (1992).
24. F. C. Chou *et al.*, *Phys. Rev. Lett.* **71**, 2323 (1993).
25. Y. Kohsaka *et al.*, *Science* **315**, 1380 (2007).
26. Y.-S. Lee, K. Segawa, Y. Ando, D. N. Basov, *Phys. Rev. B* **70**, 014518 (2004).
27. S. M. Hayden, H. A. Mook, P. Dai, T. G. Perring, F. Dogan, *Nature* **429**, 531 (2004).
28. V. Hinkov *et al.*, *Nature* **430**, 650 (2004).
29. V. Hinkov *et al.*, *Nat. Phys.* **3**, 780 (2007).
30. P. A. Marchetti, G. Orso, Z. B. Su, L. Yu, *Phys. Rev. B* **69**, 214514 (2004).
31. V. N. Kotov, O. P. Sushkov, *Phys. Rev. B* **72**, 184519 (2005).
32. M. Fujita *et al.*, *Phys. Rev. B* **65**, 064505 (2002).
33. B. Lake *et al.*, *Nature* **415**, 299 (2002).
34. N. Doiron-Leyraud *et al.*, *Nature* **447**, 565 (2007).
35. A. J. Millis, M. Norman, preprint (available at http://arxiv.org/abs/0709.0106).
36. R. Liang, D. A. Bonn, W. N. Hardy, *Phys. Rev. B* **73**, 180505 (2006).
37. We thank A. Suchaneck, S. Pailhès, and A. Amato for help during the experiments. V.H., C.B., and B.K. acknowledge financial support by the Deutsche Forschungsgemeinschaft in the consortium FOR538 and C.B. by the Schweizer Nationalfonds (SNF) via grant 200021-111690/1. We thank Y. Ando, P. Hirschfeld, G. Khaliullin, S. Kivelson, D. Manske, O. Sushkov, J. Tranquada, M. Vojta, H. Yamase, and R. Zeyher for helpful discussions.

Measuring the Surface Dynamics of Glassy Polymers

Z. Fakhraai and J. A. Forrest*

The motion of polymer chain segments cooled below the glass transition temperature slows markedly; with sufficient cooling, segmental motion becomes completely arrested. There is debate as to whether the chain segments near the free surface, or in thin films, are affected in the same way as the bulk material. By partially embedding and then removing gold nanospheres, we produced a high surface coverage of well-defined nanodeformations on a polystyrene surface; to probe the surface dynamics, we measured the time-dependent relaxation of these surface deformations as a function of temperature from 277 to 369 kelvin. Surface relaxation was observed at all temperatures, providing strong direct evidence for enhanced surface mobility relative to the bulk. The deviation from bulk α relaxation became more pronounced as the temperature was decreased below the bulk glass transition temperature. The temperature dependence of the relaxation time was much weaker than that of the bulk α relaxation of polystyrene, and the process exhibited no discernible temperature dependence between 277 and 307 kelvin.

Over the past 15 years there has been debate as to whether the molecular surface of polystyrene (PS) is glassy or not (1, 2) at temperatures below the bulk glass transition temperature. This question applies to other polymers and surface properties and has important implications for friction, lubrication, adhesion, and any applications involving polymer modification by way of coatings. Surfaces of polymers that do not crystallize, such as atactic PS, represent excellent model systems for amorphous material surfaces. There is growing evidence that the properties of the free polymer surface play a crucial role in observed anomalies in the glass transition temperature of thin polymer films (3).

Two types of experiments have been used to probe the relaxation properties of the polymer surface. One class of experiments measures changes in a physical property of the surface combined with established techniques to measure the glass transition temperature, T_g. In the case of PS, experiments using positron annihilation to directly probe the near-surface region have yielded contradictory results (4, 5), whereas experiments involving T_g measurements of a fluorescently labeled surface layer have given strong evidence for a reduced surface T_g value (6). Any measurement of a T_g probes the average response of the system at a certain relaxation time, depending on the experimental conditions (such as cooling/heating rate) and cannot provide any more information about the temperature dependence of the relaxation function. For this reason, measurements of T_g are at best indirect probes of the structural relaxation. A second class of experiments involves monitoring the response of the material to a surface perturbation. This perturbation can be oscillatory, as in the case of atomic force microscopy (AFM)–based surface rheology experiments (7, 8), or it can be applied and then allowed to relax over time (9–14). However, AFM-based methods usually use a time-dependent and often oscillatory perturbation. Recent results show that anomalous dynamics in thin films are observed only for slow (longer than 1 s) relaxation processes (15). If observed anomalies in thin-film T_g values are indeed related to the surface properties, the implication is that sub-Hz frequencies may have to be used. AFM-based techniques have been used to support (7) and refute (8) the idea of enhanced surface mobility.

For all such experiments, introduction of a surface perturbation necessarily involves introducing a stress. This stress should be as small as possible and should certainly be smaller than the yield stress of PS ($\sim 3 \times 10^7$ Pa) (16). Many experiments involving AFM tips in contact with the sample, as well as samples that are char-

Department of Physics and Astronomy and Guelph-Waterloo Physics Institute, University of Waterloo, Waterloo, Ontario N2L 3G1, Canada.

*To whom correspondence should be addressed. E-mail: jforrest@uwaterloo.ca

acterized by a deformation applied at room temperature (9, 10), may be subject to a very large stress. Experiments where the surface perturbation is applied by rubbing the polymer (9, 10) have also led to conflicting results. Some experiments that probe free-surface dynamics do so while the surface is in constant contact with a solid probe; both nanoparticle embedding and AFM-based experiments suffer from this difficulty (17, 18). In such cases it is possible to argue that the near-surface region underneath the probe is affected by both the free surface and the covering particle, and that for some length scales the free-surface effect dominates. Such an argument is not ideal, as the interaction between the polymer and the solid is not necessarily well understood (19, 20).

The length scale of the perturbation is also important. Even if there is enhanced relaxation at the surface, this region involves only fractions of polymer chains. If a mobile surface region exists, it may be thought of as being tethered to a glassy substrate. There are two consequences that arise from this: (i) Measurements really should be carried out on a single length scale if possible, and this length scale should be of the order of a few nanometers. (ii) For dynamics at length scales larger than that of any region of enhanced mobility, there may never be complete structural relaxation below the bulk T_g value. Recently Papaléo et al. (21) monitored the relaxation of nanostructures produced by ion bombardment of the surfaces of poly(methyl methacrylate) (PMMA) and found strong deviations of surface relaxation from bulk behavior. Unfortunately, these nanodeformations were produced by a very-high-energy process, and the bombardment could produce high local stress, radiation-induced cross-links, or chemical changes; these authors noted (21) that the latter two effects likely occurred within a few nanometers of the probed region.

We describe an experiment aimed at circumventing many of these experimental difficulties and provide definitive evidence for enhanced mobility at the polymer surface. We produced nanodeformations (holes) by coating a solution of gold nanospheres (diameter 23 ± 3 nm) onto the surface of a spin-cast and annealed (413 K for 12 hours under dry nitrogen) PS film (thickness 100 nm, M_w = 641,000, M_w/M_n = 1.11; Polymer Source Inc.). The nanospheres were prepared in our lab by means of a standard citrate reduction technique (22) and were thus charge-stabilized against aggregation (rather than surfactant-stabilized). We then heated the samples at 378 K for 10 to 15 min to allow the spheres to partially embed into the polymer surface (18). We chose embedding times and temperatures that produced average embedding values of ~2 to 4 nm. An embedding temperature above the bulk T_g of PS was chosen to ensure that the relaxation time of the system was less than the embedding time and that the

structure was locally relaxed around the gold nanoparticles.

After embedding the gold nanoparticles into the PS surface, it is necessary to remove them without applying a large stress to the surface. This is done by placing a drop of mercury on the surface. The mercury forms an amalgam with the gold, which essentially dissolves the gold into the mercury droplet (23). When the sample is turned at an angle, the mercury droplet slides off. After this process, what is left in place of the nanoparticles are small holes that can be measured by AFM. The imaging was done with an atomic force microscope (Explorer, Veeco) operating in tapping mode. Images of the sample after removal of the gold nanospheres in this way did not show any evidence for surface damage. (When water was used to remove the gold particles, large craters were observed in place of the gold particles, indicating surface damage due to large induced stress on the surface.) The T_g values of the films made using this procedure were measured using ellipsometry (EXACTA 2000, Waterloo Digital Electronics) to be the same as the bulk value, and energy-dispersive x-ray analysis (LEO 1350 SEM system) did not show residual mercury or gold on the polymer surface.

Figure 1 illustrates the sample preparation technique. In comparing Fig. 1A and Fig. 1B, it is clear that in place of nanospheres there are holes surrounded by a rim of PS. Figure 1C shows a histogram of hole depth data (from a 5 µm by 5 µm image) from which we determine the average hole depth. A line profile used to find the depth of a single hole is shown in Fig. 1D. Each sample is measured after nanohole formation to ensure that there are enough holes on the surface to get adequate statistics. There is an inherent sensitivity to the sharpness of the AFM tip; if the radius of curvature of the AFM tip is much larger than that of the nanohole, there is a reduced ability to measure the depth of the hole. This problem is minimized by using sharp tips and frequently changing the tip. The average initial depth of the nanodeformations can be compared to the average embedding depth of the spheres before dissolution to ensure that the tips and the imaging resolutions are chosen properly and the depths of the nanodeformations can be measured accurately and reproducibly. This technique is applicable to other polymers, and we have also seen similar effects in isotactic PMMA.

The principal driving force that relaxes the surface roughness is stress due to the polymer surface tension. The magnitude of the initial stress σ_0 depends on the radius of curvature of the nanodeformation: $\sigma_0 = 2\gamma/R$, where γ is the PS surface tension and R is the radius of the nanosphere used to make the hole. For a spherical cap with a radius of curvature of 11.5 nm, $\sigma_0 = 1.3 \times 10^6$ Pa. This is at least an order of magnitude smaller than the yield stress of PS and provides confidence that we are observing a linear response. The stress gradually decreases as the nanohole anneals and becomes flat. This is the
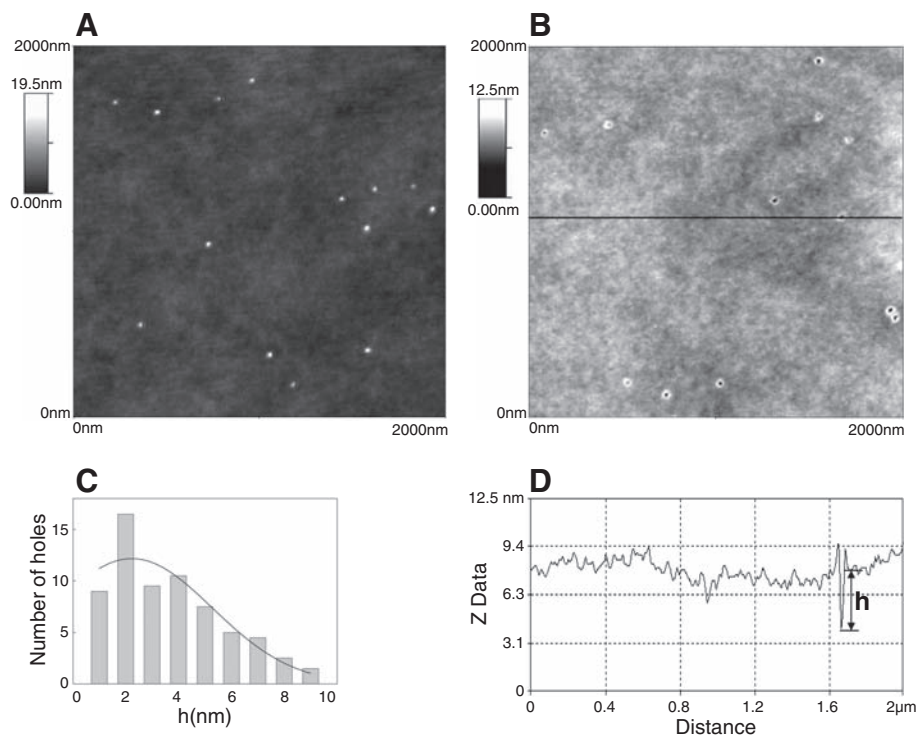
Fig. 1. Illustration of the process used to create nanoholes in the PS surface. (**A**) PS surface covered by Au nanospheres that have been partially embedded. (**B**) The surface of the same PS sample, but at a different location, after Au nanospheres have been removed by exposure to liquid Hg. (**C**) Histogram showing hole depth. (**D**) Line scan of the indicated line in (**B**) used to measure the depth of a single hole. Note the existence of the rim around the hole.

only driving force present in the system that can cause a complete annealing of the surface deformations. The stress on the sample due to a mismatch in thermal expansion between the polymer and the gold particles results in at most a 1% change in the nanodeformation and would not be observable. The depth of the surface region probed is determined by the extent to which the particles are allowed to embed, and the initial driving force for annealing at the bottom of the nanodeformation is determined only by R and is independent of the embedding depth, h. The embedding depth defines the lateral radius of the hole to be $[R^2 - (R-H)^2]^{1/2}$. These two quantities, R and h, can be independently varied (subject to the constraint $h < R$) by the choice of sphere size and annealing time. For shallow holes like those used in our experiment, it is only the measurement of the hole radius, and not the depth, that is limited by the shape of the AFM tip.

To avoid extra stress caused by repeated heating and cooling of the samples, we measure each sample only once after annealing has begun. This means that we need at least one sample per annealing time, and determination of the relaxation curve for any sample temperature requires 10 to 15 samples with identical thermal history and similar initial hole depth distributions. For a given sample temperature, all of the prepared samples are placed in the sample oven in dry nitrogen. After some time interval, one of the samples is removed, cooled down to room temperature in a few seconds, and then measured by AFM. Each AFM image contains many holes (~50), and we use the average depth over all holes. This procedure is repeated at a number of elapsed annealing times until holes are no longer observed. Even when the holes have relaxed to the point where their depth is too small to be measured, their rims persist and so they can still be located; the depths of these holes are assumed to be zero. If these holes simply disappeared, then we would only observe the largest holes at long times and the resulting analysis would provide average values that are too large. The measured hole depths are used to determine the time dependence of the indentation depth at constant annealing temperature. The entire experiment is then performed at different annealing temperatures.

Figure 2 shows the time evolution of nanodeformations at a sample temperature of 293 K. As the annealing time is increased, both the depth of the holes and the size of the rims are decreased. The rims relax more slowly because they have a different shape with a larger radius of curvature (and hence a smaller driving force) than the holes. The holes (and rims) relax at a temperature that is almost 80 K below the bulk T_g value. Images such as those in Fig. 2, but over a larger distance (typically 5 μm by 5 μm), are used to provide the hole depth values for each time and annealing temperature. At all measured temperatures, the time dependence of

the average hole depths is well described by a single-exponential function within the scatter of the data. Hence, it is possible to find a relationship between the time constant of this annealing and the characteristic relaxation time of the system near the free surface. The depth of a hole as a function of time is described by $h = h_0 \exp(-t/\tau)$, where τ is the time constant of the annealing. Because at each given time the surface stress is derived from the radius of curvature of the deformation at that time, it can be assumed that at the bottom of the hole, the stress due to the surface tension also exhibits single-exponential decay with the same time constant τ. Because the time dependences of both the stress and strain functions are single-exponential, the time and position dependences can be separated

$$\varepsilon_{ij}(x_i, t) = \varepsilon_{ij}(x_i) \exp(-t/\tau) \qquad (1)$$

$$\sigma_{ij}(x_i, t) = \sigma_{ij}(x_i) \exp(-t/\tau) \qquad (2)$$

where ε and σ are the general strain and stress functions. Given the single-exponential relaxation observed, it is a reasonable assumption that the creep modulus has a similar dependence

$$G = G_0 \exp(-t/\tau_\alpha) \qquad (3)$$

where G_0 is the longitudinal glassy creep modulus, and τ_α is the characteristic time that is

governing the dynamics of the observed surface relaxation. Separating the position dependence reveals the time-dependent part of the differential equation of the system (24)

$$\frac{1}{\tau_\alpha}\sigma(t) + \frac{d\sigma(t)}{dt} = G_0\varepsilon(t) \qquad (4)$$

Using this equation, it is easy to show that for single-exponential relaxation

$$\tau = \tau_\alpha\left(1 + \frac{RG_0}{2\gamma}\right) \qquad (5)$$

The longitudinal modulus is used because we are only measuring changes in the depth of the hole, and only a longitudinal stress drives the changes in depth at the bottom of the nanohole. This relationship can be used to obtain the characteristic relaxation time of the surface from the measured hole annealing time constants.

The time evolution of the hole depth can be used to obtain the surface relaxation function for each temperature studied. If the shape of the relaxation function is not temperature-dependent, then it is possible to find a shift factor, α_T, in time such that the relaxation curves at different temperatures can be superimposed. This is a commonly used technique to describe the temperature dependence of relaxation times in glass-forming materials. Figure 3 shows the superposition plot as well as the shift factors used to arrive at the
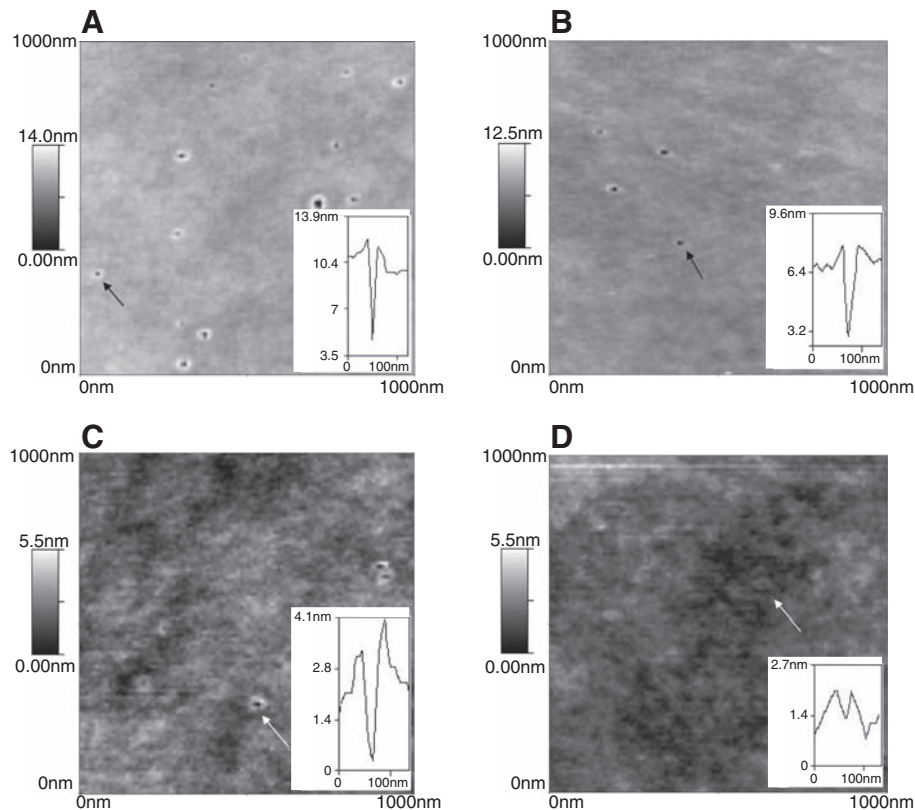
Fig. 2. Evolution of nanoholes at a sample temperature of 293 K. Annealing times are (**A**) 0 hours, (**B**) 74 hours, (**C**) 495 hours, and (**D**) 1607 hours. The inset of each panel shows the line scan of the indicated holes used to find their depth.

cumulative plot. Note that each data point on this plot is obtained from a different sample, and the resulting scatter is mainly a result of the fact that each sample has a slightly different distribution of hole sizes (all with average values between 2 and 4 nm). The data in Fig. 3 are scaled to a reference temperature of 369 K, which is the highest temperature for which we were able to get reliable relaxation measurements. It is clear that the relaxation function can be well described by a single exponential. This is in contrast to the relaxation function of bulk PS, which exhibits stretching ($\phi = \phi_0 \exp[-(t/\tau)^\beta]$) with a β value of 0.4 (dashed curve in Fig. 3).

Large deviations from the bulk are also seen in the shift factors used to collapse the relaxation functions. For bulk PS at temperatures greater than T_g, the shift factors for PS have the form $\log(a_T) = [C_1(T - T_{ref})]/(C_2 + T - T_{ref})$, where $C_1 = 12.7$ to 13.7, $T_{ref} = 373$ K, and $C_2 = 49.9$ (25) [solid ($T > T_g$) and dotted ($T < T_g$) curves, inset of Fig. 3]. An obvious consequence of this form is the divergence of shift factors at a temperature $T^* = T_{ref} - C_2 = 322.9$ K. Below bulk T_g, the sample is in a nonequilibrium glassy state. In this state the measured dynamics are faster than an extrapolation of the liquid dynamics and take on an Arrhenius form (solid line, inset of Fig. 3). This

enhanced α relaxation in the glassy state (aging) is still substantially slower than what we observed for the surface; more important, it has a different temperature dependence. For the PS surface relaxation, at lower temperatures the temperature dependence of the shift factors becomes so weak that no substantial temperature dependence is observed below 307 K. In the inset of Fig. 3, the bulk dynamics are obtained from the α relaxation measurements as probed by the relaxation of dye molecules using the second harmonic generation (SHG) technique (26). The surface relaxation data clearly do not obey the behavior of bulk PS; instead, the slope decreases with decreasing temperature. Figure 3 represents a reasonably complete quantification of the dynamics of the top 2- to 4-nm surface layer of PS below the bulk glass transition temperature.

A more detailed comparison to the bulk α process (the main structural relaxation process in bulk glass-formers) is possible. Equation 5 can be used to convert the shift factors into temperature-dependent relaxation times of the surface. Figure 4 shows the relaxation times along with plots of α and β (local vibrational relaxation) of PS. Note that because different techniques are used for the bulk (26) and surface measurements, there may be a relative vertical shift (about one or two

orders of magnitude) in one or both of the data sets, but the temperature dependences are independent of the technique and can be used to compare the two data sets. It can be seen that the relaxation times of the surface are similar to the bulk α relaxation times near the bulk value of T_g but deviate strongly from the bulk as the temperature is decreased further below the bulk value of T_g. To address the idea of using different driving forces (such as one would get with different sphere sizes), this plot also shows an analysis of the rim data (solid symbols) where they could also be determined. The similarity of the hole and rim data, combined with the fact that the radius of curvature of the rims is about twice as large (and hence subject to half the stress) as that of the holes, provides more evidence that the observed relaxation is a linear viscoelastic response.

Although the analysis provided by Eqs. 1 to 5 provides for a definite scale between the observed relaxation times and the characteristic relaxation time of the near-surface region, the essential results of this work (definitive enhanced surface mobility and weak temperature dependence of relaxation times) are robust and model-independent. Note that for temperatures near the bulk T_g, the relaxation time of the surface is very similar to the bulk α relaxation time, as was also observed in (21) for PMMA. As the temperature is decreased below the bulk T_g, the bulk dynamics freeze rapidly while the surface relaxation time changes only about two orders of magnitude over a temperature range of about 80 K. This results in a growing disparity between surface and bulk properties as the temperature is lowered below T_g. Of particular interest is that the temperature dependence of the surface relaxation becomes weaker as the temperature is decreased (opposite to the temperature dependence typical of glass-forming behavior, where an Arrhenius temperature dependence with a constant activation energy exists below T_g), and between 277 and 307 K no temperature dependence was discerned. These results alone are not sufficient to determine whether the surface relaxation is an α process that strongly deviates from the bulk, or whether it is a previously unobserved mode of relaxation that is only available near the free surface. The temperature dependence of the surface relaxation process can provide an explanation of some apparent contradictions in the literature. For example, experiments that probe the dynamics above or near the bulk T_g would show that the dynamics of the surface are very similar to the bulk dynamics—a result that would contradict measurements made at lower temperatures.

Our experiment enabled us to measure the temperature-dependent relaxation time for the very near (2 to 4 nm) surface region of PS films. The surface region showed relaxation at all temperatures measured (down to 277 K), and the resulting relaxation times displayed a surprisingly weak temperature dependence below the bulk T_g value, which became immeasurable for temperatures less than ~307 K. These results are

Fig. 3. Time dependence of the relaxation of nanoholes for all temperatures. The time data are shifted by a factor a_T so that they agree with the data at the reference temperature of 369 K. The solid line is a single-exponential fit; the dashed line is a stretched exponential with bulk value of $\beta = 0.4$, with the same average lifetime as the single-exponential fit. The inset shows logarithms of the shift factors used to produce superposition; the solid curve in the inset corresponds to the bulk shift factors of PS above and below bulk T_g, obtained by the SHG technique (26).
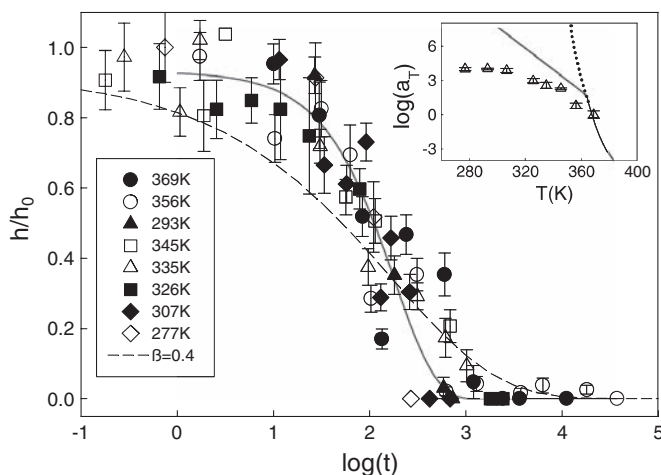
Fig. 4. Comparison of the calculated surface relaxation times (open triangles) with bulk α (26) and β (27) relaxations of PS. The solid curve shows α relaxation times in PS obtained from the SHG technique (26). The dashed curve indicates the β relaxation times of PS obtained from dynamic mechanical measurements (27). The solid triangles denote relaxation times obtained from the annealing of the rims, which have a larger radius of curvature than the holes.
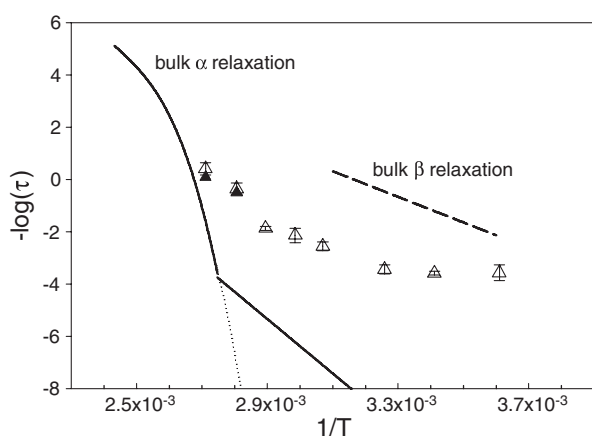
in stark contrast to the regular α and β relaxation processes that are observed in bulk glass-formers.

References and Notes

1. G. F. Meyers, B. M. DeKoven, J. T. Seitz, *Langmuir* **8**, 2330 (1992).
2. J. A. Forrest, *Eur. Phys. J. E* **8**, 261 (2002).
3. J. S. Sharp, J. A. Forrest, *Phys. Rev. Lett.* **91**, 235701 (2003).
4. L. Xie *et al.*, *Phys. Rev. Lett.* **74**, 4947 (1995).
5. Y. C. Jean *et al.*, *Phys. Rev. B* **56**, R8459 (1997).
6. C. J. Ellison, J. M. Torkelson, *Nat. Mater.* **2**, 695 (2003).
7. J. Hammerschmidt, W. Gladfelter, G. Haugstad, *Macromolecules* **32**, 3360 (1999).
8. S. Ge *et al.*, *Phys. Rev. Lett.* **85**, 2340 (2000).
9. Y. Liu *et al.*, *Macromolecules* **30**, 7768 (1997).
10. A. D. Schwab, A. Dhinojwala, *Phys. Rev. E* **67**, 021802 (2003).
11. T. Kerle, Z. Lin, H. Kim, T. P. Russell, *Macromolecules* **34**, 3484 (2001).
12. E. Buck, K. Petersen, M. Hund, G. Krausch, D. Johannsmann, *Macromolecules* **37**, 8647 (2004).
13. P. Gasemjit, D. Johannsmann, *J. Polym. Sci. B* **44**, 3031 (2006).
14. I. Karapanagiotis, W. W. Gerberich, *Macromolecules* **38**, 3420 (2005).
15. Z. Fakhraai, J. A. Forrest, *Phys. Rev. Lett.* **95**, 025701 (2005).
16. G. Strobl, *The Physics of Polymers* (Springer-Verlag, Berlin, ed. 2, 1997).
17. V. M. Rudoy *et al.*, *Colloid J.* **64**, 746 (2002).
18. J. H. Teichroeb, J. A. Forrest, *Phys. Rev. Lett.* **91**, 016104 (2003).
19. J. S. Sharp *et al.*, *Eur. Phys. J. E* **22**, 287 (2007).
20. S. A. Hutcheson, G. B. McKenna, *Phys. Rev. Lett.* **94**, 076103 (2005).
21. R. M. Papaléo *et al.*, *Phys. Rev. B* **74**, 094203 (2006).
22. J. Turkevich, G. Garton, P. C. Stevenson, *J. Colloid Sci.* **9**, 26 (1954).
23. This technique was first published in 1556 by Georgius Agricola in *De Re Metallica*, but it may have been used even in Roman times.
24. R. M. Christensen, *Theory of Viscoelasticity* (Academic Press, New York, 1971).
25. J. D. Ferry, *Viscoelastic Properties of Polymers* (Wiley, New York, ed. 3, 1980).
26. A. Dhinojwala, G. K. Wong, J. M. Torkelson, *J. Chem. Phys.* **100**, 6046 (1994).
27. J. Y. Cavaille, C. Jourdan, J. Perez, *J. Polym. Sci. B* **25**, 1235 (1987).
28. We thank S. Valadkhan and M. Stastna for helpful discussions, and the Natural Sciences and Engineering Research Council (Canada) for financial support of this project.

Abiogenic Hydrocarbon Production at Lost City Hydrothermal Field

Giora Proskurowski,[1,2]* Marvin D. Lilley,[1] Jeffery S. Seewald,[2] Gretchen L. Früh-Green,[3] Eric J. Olson,[1] John E. Lupton,[4] Sean P. Sylva,[2] Deborah S. Kelley[1]

Low-molecular-weight hydrocarbons in natural hydrothermal fluids have been attributed to abiogenic production by Fischer-Tropsch type (FTT) reactions, although clear evidence for such a process has been elusive. Here, we present concentration, and stable and radiocarbon isotope, data from hydrocarbons dissolved in hydrogen-rich fluids venting at the ultramafic-hosted Lost City Hydrothermal Field. A distinct "inverse" trend in the stable carbon and hydrogen isotopic composition of C_1 to C_4 hydrocarbons is compatible with FTT genesis. Radiocarbon evidence rules out seawater bicarbonate as the carbon source for FTT reactions, suggesting that a mantle-derived inorganic carbon source is leached from the host rocks. Our findings illustrate that the abiotic synthesis of hydrocarbons in nature may occur in the presence of ultramafic rocks, water, and moderate amounts of heat.

Fischer-Tropsch type (FTT) reactions involve the surface-catalyzed reduction of oxidized carbon to CH_4 and low-molecular-weight hydrocarbons under conditions of excess H_2. This set of reactions has been commonly invoked to explain elevated hydrocarbon concentrations in hydrothermal fluids venting from submarine ultramafic-hosted systems (*1*) and in springs issuing from ophiolites (*2*); however, whether naturally occurring FTT reactions are an important source of hydrocarbons to the biosphere remains unclear. Although CH_4 and higher hydrocarbons have been synthesized by FTT in the gas phase from CO for more than 100 years (*3*), only recently were FTT reactions shown to proceed, albeit with low yields, under aqueous hydrothermal conditions, with dissolved CO_2 as the carbon source (*4*, *5*). The reactions involved in Fischer-Tropsch reduction of aqueous CO_2 can be expressed in general terms by the reaction

$$CO_{2aq} + [2 + (m/2n)]H_2 \rightarrow (1/n)C_nH_m + 2H_2O \quad (1)$$
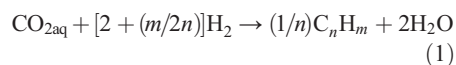
Here, we show that low-molecular-weight hydrocarbons in high-pH vent fluids from the ultramafic-hosted Lost City Hydrothermal Field (LCHF) at 30°N on the Mid-Atlantic Ridge (MAR) are likely produced abiotically through FTT reactions.

The LCHF is situated near the summit of the Atlantis Massif, ~15 km west of the MAR axis. Towering carbonate chimneys (up to 60 m tall) diffusely vent high-pH (9 to 11), moderate-temperature (28° to 90°C) fluids, produced by reaction of seawater with rocks originating from the mantle (*6*, *7*). The basement directly beneath this system consists of highly serpentinized peridotites (dominated by depleted mantle harzburgites) with lesser talc schists and metagabbros exposed by long-lived detachment faulting (*8*, *9*). Fluid circulation is driven by cooling of the underlying rocks (*10*), perhaps supported by minor contributions from exothermic serpentinization reactions (*7*, *11*) or a yet undetected magmatic source.

Serpentinization is the hydration of olivine and orthopyroxene minerals, the main constituents of ultramafic rocks, and creates a reducing chemical environment characterized by high H_2 concentrations that is well suited to abiotic hydrocarbon production. The general reaction is

$$6[(Mg_{1.5}Fe_{0.5})SiO_4] + 7H_2O \rightarrow$$
$$3[(Mg_3Si_2O_5(OH)_4] + Fe_3O_4 + H_2 \quad (2)$$

(olivine, water → serpentine, magnetite, hydrogen)
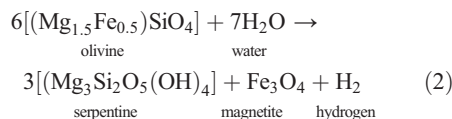
At Lost City, vent fluids have end-member H_2 concentrations of 0.5 to 14.4 mmol/kg (*12*), greater than the highest H_2 concentrations in fluids sampled within basaltic-hosted environments that are unperturbed by magmatic and eruptive events (*6*, *13*). The highest H_2 concentrations within the LCHF approach those measured at Rainbow (16 mmol/kg) and surpass those measured at Logatchev (12 mmol/kg); both are peridotite-and-gabbro–hosted systems on the MAR (*1*). Rainbow and Logatchev vent fluids differ from those at the LCHF in that they have elevated concentrations of CO_2 and dissolved metals, low pH, and substantially higher temperatures (350° to 360°C): characteristics typical of high-temperature black smoker vents hosted in mid-ocean ridge basalts (MORBs) (*1*). In contrast, the moderate-temperature, high-pH, sulfate-bearing fluids at the LCHF are enriched in H_2, CH_4, and low-molecular-weight volatile hydrocarbons, but are highly depleted in CO_2 and dissolved metal contents.

The chemistry of LCHF fluids implies that fluid temperatures beneath the seafloor may be 200° ± 50°C (*10*), although oxygen isotope data from carbonate veins and low D/H ratios of H_2 argue for temperatures <150°C (*8*, *12*). In general, H_2 concentrations in LCHF fluids are lower than values observed during laboratory experiments (*14*) and predicted by equilibrium models (*10*) and likely reflect a more limited extent of serpentinization at the lower temperatures associated with the natural system (*15*). The low and variable H_2 concentrations in the LCHF fluids may also reflect some removal and use of H_2 by

[1]School of Oceanography, University of Washington, Seattle, WA 98195, USA. [2]Woods Hole Oceanographic Institution, Woods Hole, MA 02543, USA. [3]Department of Earth Sciences, ETH-Zentrum, Zurich, Switzerland. [4]National Oceanic and Atmospheric Administration (NOAA)–Pacific Marine Environmental Laboratory, Newport, OR 97365, USA.

*To whom correspondence should be addressed. E-mail: giora@whoi.edu

Table 1. Carbon and hydrogen isotope data from Lost City vents. All isotope values are in ‰ units; $\delta^{13}C$ is reported as vPDB, and δD as vSMOW.

Sample	Marker	$\delta^{13}C\text{-}CH_4$*	$\delta D\text{-}CH_4$†	$\delta^{13}C\text{-}C_2H_6$‡	$\delta D\text{-}C_2H_6$§	$\delta^{13}C\text{-}C_3H_8$‡	$\delta D\text{-}C_3H_8$§	$\delta^{13}C\text{-}n\text{-}C_4H_{10}$‡
3879-GT10	H	−13.6		−14.9				
3881-GT10	H	−13.2		−14.3				
H04-IGT8	H	−9.5	−119	−15.2	−127	−15.0	−140	
3865-GT7	IF	−12.3	−139	−13.1		−13.9		−13.0
H05-IGT7	IF	−11.8	−147	−13.7	−166	−13.4	−160	
3871-GT15	8	−9.9	−136	−13.3		−14.2		−14.3
3866-GT14	7	−13.2	−125	−14.6		−15.8		−14.9
3871-GT9	C	−12.1	−126	−13.5		−14.2		−13.8
3876-GT7	BH	−11.0	−127	−13.5		−14.5		−14.6
H06-IGT8	BH	−9.4	−131	−14.5	−152	−13.8	−133	
3863-GT7	3	−10.5		−13.9				
H03-IGT7	3	−11.8	−131	−14.3	−151	−14.0	−171	
3872-GT10	6	−11.8		−14.4		−16.0		
H07-IGT7	6	−12.0	−135	−14.9	−138			

*$\delta^{13}C\text{-}CH_4$ measurement uncertainty ± 0.2‰. †$\delta D\text{-}CH_4$ measurement uncertainty for 2005 samples (H0x-IGTx) ± 1.1‰, for 2003 samples (38xx-GTx) ± 6.5‰. ‡$\delta^{13}C\text{-}C_{2\text{-}4}$ measurement uncertainty ± 0.3‰. §$\delta D\text{-}C_{2\text{-}3}$ measurement uncertainty ± 4.1‰.
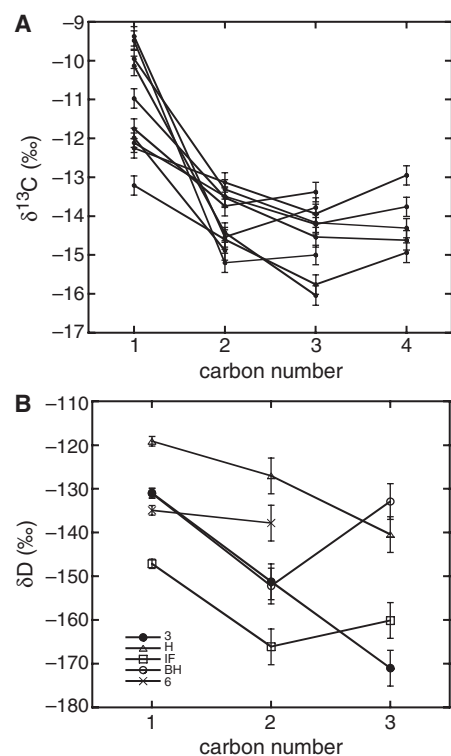
Fig. 1. ^{13}C and D trends. (**A**) $\delta^{13}C$ versus carbon number for C_1 to C_4 LCHF alkanes. (**B**) δD versus carbon number for C_1 to C_3 LCHF alkanes. Trend of increasing isotopic depletion with increasing carbon number is opposite to the observed isotopic trend for thermogenically produced alkanes.

microbes. The porous carbonate structures of active vents at Lost City, created by the mixing of vent fluid with seawater, have high microbial cell counts, and phylogenetic studies indicate the presence of H_2-utilizing microbes (7, 16, 17). Additionally, 16S ribosomal RNA genes corresponding to those of methanogens, aerobic and anaerobic methanotrophs, sulfate reducers, and sulfur oxidizers were detected in these energy-rich environments (16).

Lost City fluids have CH_4 concentrations (1 to 2 mmol/kg) greater than values from un-sedimented basalt-hosted hydrothermal systems, but similar to values in fluids from the serpentine-hosted Rainbow and Logatchev vent fields (1, 16). The concentrations are low relative to those of hydrothermal fluids from sediment-hosted environments, where CH_4 is produced by the thermal decomposition of organic matter within the sediment (16, 17). Thermogenic CH_4 in such sedimented systems is commonly characterized by $\delta^{13}C$ values of −50 to −30 per mil (‰) (18, 19). Methane $\delta^{13}C$ values at LCHF range from −14 to −9‰ [all carbon and hydrogen isotopic values are given in standard δ notation and ‰ units, referenced to the Vienna Pee Dee belemnite (vPDB) and Vienna standard mean ocean water (vSMOW) standards, respectively]. Lost City $\delta^{13}C$ values of CH_4 are similar to values of CH_4 hypothesized to be of abiogenic origin from Rainbow (−16‰), Logatchev (−14‰), and the ultramafic-hosted Zambales ophiolite seeps (−7‰) (1, 16). The high $\delta^{13}C$ values of CH_4 from Lost City fluids, as well as the lack of a sediment source rich in organic matter along the reaction path, suggest that CH_4 is not appreciably derived from a thermogenic source.

The isotopic composition of short-chain hydrocarbons at LCHF suggests that abiotic synthesis is responsible for their formation. The carbon isotope compositions of C_1 to C_4 hydrocarbons from LCHF fluids are increasingly negative ($\delta^{13}C$ ranges from −9 to −16‰) with increasing chain length (Table 1 and Fig. 1A). This isotopic pattern is opposite to that for hydrocarbons produced thermogenically (20). Such an "inverse" isotopic trend ($\delta^{13}C_1 > \delta^{13}C_2 > \delta^{13}C_3 > \dots$) has been shown experimentally to be a possible indicator of abiotic synthesis, and specifically of FTT reactions under conditions of incomplete (35 to 75%) conversion of CO (21). Abiotic synthesis has also been invoked to account for decreases in $\delta^{13}C$ values of C_2 to C_4 alkanes with respect to CH_4 in other natural systems (22). The hydrogen isotopic composition of Lost City C_1 to C_3 hydrocarbons

shows a similar, although less defined, trend in which molecules of longer chain length have similar or slightly lower δD values (−120 to −170‰) relative to shorter-chain alkanes (Fig. 1B). The pattern is, again, opposite to the trend expected for thermogenic gases (23), and also opposite to the trends for gases from Precambrian rocks in the Canadian Shield where an enrichment in D with increasing chain length (−425 to −250‰) was attributed to formation by polymerization of CH_4 (22). The differences between LCHF and the Canadian Shield D/H trends likely reflect different formation mechanisms and require further investigation through laboratory experiments.

In addition to their isotopic composition, the abundance and relative distribution of short-chain hydrocarbons are consistent with an abiotic formation mechanism. At Lost City, the alkanes and ethene make up most of the volatile-gas composition, although acetylene, propene, and propyne are present in several samples (table S1). A plot of the log of the n-alkane concentration against carbon number shows a strong linear correlation with carbon number for C_2 to C_4 alkanes, with elevated CH_4 concentrations (fig. S1D). This trend is consistent with an Anderson-Schulz-Flory distribution predicted for FTT synthesis (3). However, such a trend is also consistent with thermogenic production of low-molecular-weight n-alkanes (24).

The elevated CH_4 concentrations, relative to those of C_2 to C_4 n-alkanes, suggests that FTT synthesis may not reflect the only abiotic source of hydrocarbons to LCHF fluids. Laboratory experiments have demonstrated that reduction of CO_2 to CH_4 under hydrothermal conditions in the presence of Fe-Ni alloys is rapid and does not result in the production of C_{2+} hydrocarbons (25). Ultramafic rocks from the Atlantis Massif have Ni concentrations of 1700 to 2400 parts per million (ppm), a range that is representative of dredged ultramafic samples from numerous locations along the MAR (26). Although most of the Ni in ultramafic rocks is incorporated into silicate minerals and not catalytically available, Fe-Ni

alloys are viable catalysts and stable under highly reducing conditions associated with early stages of serpentinization and lower water/rock ratios (W/Rs) (27).

Radiocarbon measurements represent a powerful tool to constrain the origin of carbon involved in abiotic synthesis of hydrocarbons. Accelerator mass spectrometric measurements on six aqueous CH_4 samples from LCHF consistently yielded ^{14}C contents near the detectable limit (Table 1), i.e., the CH_4 at LCHF is radiocarbon dead (28). Thus, the carbon source of CH_4 cannot be seawater bicarbonate (^{14}C-seawater$_{DIC}$ measured as modern) that has been microbially or abiogenically reduced. The absence of a modern seawater bicarbonate signal requires that bicarbonate be removed before production of the hydrocarbons in the vent fluids. This likely occurs by precipitation as $CaCO_3$, either in the more permeable zones of seawater recharge, as observed in the recharge limbs of circulation cells at black smoker systems (29) and in ridge-flank environments (30), or in the serpentinizing basement, where increasing pH leads to carbonate precipitation within the serpentinites (8, 10).

The ^{14}C content of short-chain hydrocarbons suggests that the requisite carbon for abiotic synthesis is derived by leaching of primordial radiocarbon-dead carbon from mantle host rocks. Mantle rocks and hydrothermal vent fluids typically have CO_2/[3]He ratios of $\sim 1 \times 10^9$ (31). In contrast, the ratios in the LCHF fluids are much lower, ranging from 3×10^4 to 1×10^7. Lost City [4]He concentrations are one order of magnitude less than those of MOR samples, and CO_2 concentrations (0.1 to 26 μmol/kg) are up to six orders of magnitude less (Fig. 2 and table S2). The low CO_2/[3]He ratios at Lost City thus primarily reflect extremely low CO_2 concentrations, suggesting that mantle CO_2 has been removed from the fluids before venting. If a standard MOR value for CO_2/[3]He is assumed, an average CO_2 concentration of 2.9 ± 0.7 mmol/kg is predicted from the He content of the fluids (table S2). On an individual sample basis, total hydrocarbon concentrations can account for 35 to 56% of predicted CO_2 concentrations. These yields are high relative to the 1% yields typical of aqueous experimental studies (4, 5), but are consistent with the 35 to 75% yields of the FTT experiments that resulted in "inverse" carbon isotope trends (21). Even higher yields are predicted if, before abiotic reduction, mantle CO_2 were removed from the system by carbonate precipitation under alkaline conditions created by serpentinization reactions.

Near-quantitative reduction of mantle CO_2 to hydrocarbons is consistent with the similar ^{13}C content of CO_2 and hydrocarbons at the LCHF. The $\delta^{13}C$ value of CO_2 at LCHF is estimated to be −9‰ (32). This value is within the −10 to −5‰ range observed for $\delta^{13}C$ of CO_2 from olivine gabbro norites from the South West Indian Ridge (33). Although laboratory experiments conducted at 250°C indicate that fractionation between CO_2 and CH_4 is ~35‰ during low-yield FTT synthesis (5) and equilibrium fractionation at this temperature is of similar magnitude, quantitative conversion of CO_2 to hydrocarbons will result in the reduced carbon species inheriting the isotopic composition of the initial CO_2. Thus, $\delta^{13}C$ values for CH_4 that range from −13.6 to −9.4‰ in LCHF fluids are consistent with the near-complete reduction of mantle-derived CO_2 with a $\delta^{13}C$ value of −9‰. Mass balance constraints (assuming a closed system, and a fractionation factor ε = 35‰) suggest that a −14‰ $\delta^{13}C$ value of CH_4 can be explained by an 85% conversion of a −9‰ CO_2 source. Extensive reduction of mantle-derived CO_2 is supported by a positive correlation between $\delta^{13}C$ values of CH_4 and percent conversion of mantle CO_2 calculated from [3]He content of the LCHF fluids (Fig. 3).

The possibility exists that microbial metabolic activity is partially responsible for the near-quantitative conversion of mantle-derived CO_2 to CH_4. Indeed, the presence of methanogens is indicated in the porous carbonate structures formed by diffuse flow (7, 34, 35). Our measurements do not counterindicate methanogenesis as a source of CH_4 to Lost City fluids; however, microbial processes cannot explain the high concentrations and distinctive isotopic signature of C_{2+} hydrocarbons.

Carbon in mantle rocks occurs in a variety of forms: primarily as inorganic carbon dissolved within the mineral matrix or trapped in fluid inclusions as graphite, or amorphous carbon residing along mineral grain boundaries (36). During hydrothermal circulation at mid-ocean ridges, there is a net transfer of carbon from the host rocks to the circulating fluid during fluid-rock reactions or through magmatic degassing. If all hydrocarbons in LCHF fluids are derived from CO_2 originally stored in the underlying rocks (or CO_2 produced from carbon in them), the W/R describing the fluid reaction history can be constrained.

Fig. 2. Measured [4]He and CO_2 concentrations. LCHF [4]He and CO_2 concentrations are depleted relative to those of other hydrothermal vent sites (table S2). The plotted line represents CO_2/[3]He = 1×10^9 (assuming a constant [3]He/[4]He, where R/Ra = 8.7), a value that typifies MORB glasses and MOR hydrothermal fluids (31). LCHF data plot well to the left of the line, indicating a loss of CO_2 at LCHF.
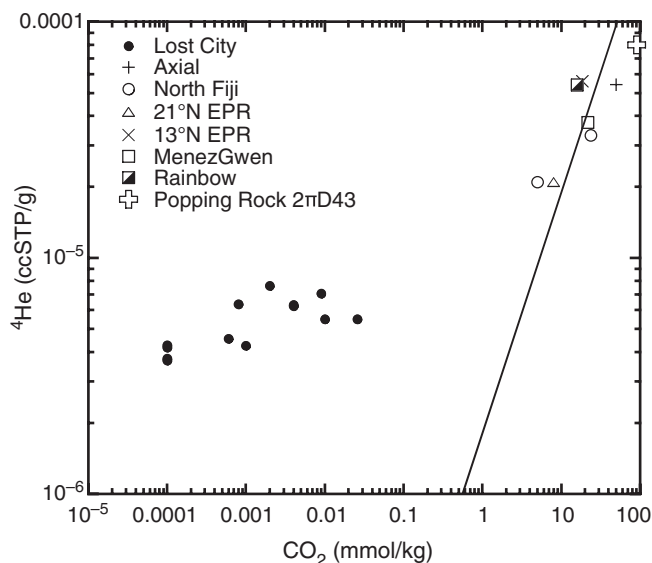
Fig. 3. [13]C composition of CH_4 varies with modeled conversion efficiency. Percent conversion is based on predicted CO_2 concentrations (as calculated from the LCHF [3]He value and a standard MOR value for CO_2/[3]He). Higher yields correspond with $\delta^{13}C$ values of CH_4 closer to the approximately −9‰ $\delta^{13}C$ values of CO_2 at Lost City. Although calculated yields are below quantitative levels, if CO_2 is removed by non–CH_4-forming mechanisms before FTT reactions (e.g., precipitation of $CaCO_3$ at high-pH conditions), these yields would increase.

y = −16.717 + 10.416x   R= 0.58626

REPORTS
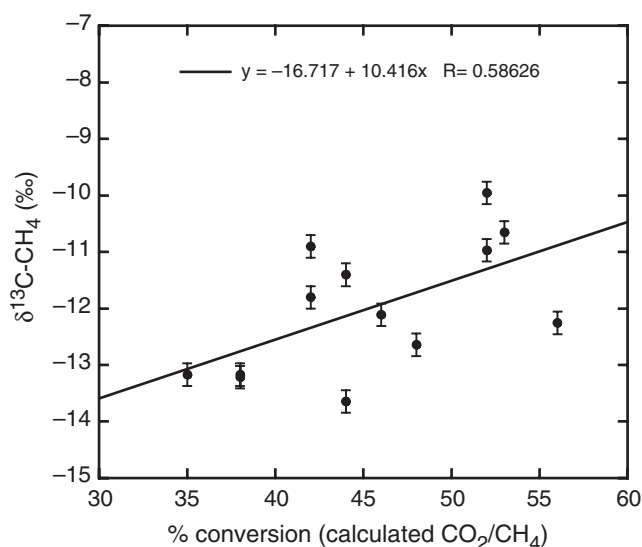
For hydrocarbon (simplified in this model to just CH_4) concentrations of 1.4 mmol/kg, the maximum W/R is 64, assuming 100% conversion of CO_2 to CH_4 and an initial (but high) CO_2 concentration of ~4000 ppm in the basement rocks (37) (fig. S2). This W/R is at the low end of those predicted from the Sr and Nd isotopic compositions of LCHF serpentinites (37); however, the samples from seafloor outcrops almost certainly have a reaction history different from that of the rocks directly supplying the present-day fluids at Lost City. More typical and lower initial basement rock CO_2 concentrations would yield lower W/Rs. On the basis of a system constrained by a 400-ppm CO_2 concentration in the basement rocks (27) and a conversion of ~50% (as suggested by the He and CO_2 data), we posit that the fluids feeding the LCHF have reacted with rocks in a W/R of less than 5 (fig. S2).

Lost City may be just one of many, as yet undiscovered, off-axis hydrothermal systems. Hydrocarbon production by FTT could be a common means for producing precursors of life-essential building blocks in ocean-floor environments or wherever warm ultramafic rocks are in contact with water.

References and Notes

1. J. L. Charlou, J. P. Donval, Y. Fouquet, B. P. Jean, N. Holm, *Chem. Geol.* **191**, 345 (2002).
2. T. A. Abrajano *et al.*, *Chem. Geol.* **71**, 211 (1988).
3. R. B. Anderson, *The Fischer-Tropsch Synthesis* (Academic Press, Orlando, FL, 1984).
4. D. I. Foustoukos, W. E. Seyfried Jr., *Science* **304**, 1002 (2004).
5. T. M. McCollom, J. S. Seewald, *Earth Planet. Sci. Lett.* **243**, 74 (2006).
6. D. S. Kelley *et al.*, *Nature* **412**, 145 (2001).
7. D. S. Kelley *et al.*, *Science* **307**, 1428 (2005).
8. G. L. Früh-Green *et al.*, *Science* **301**, 495 (2003).
9. J. A. Karson *et al.*, *Geochem. Geophys. Geosyst.* **7**, Q06016 (2006).
10. D. E. Allen, J. Seyfried, W. E. *Geochim. Cosmochim. Acta* **68**, 1347 (2004).
11. R. P. Lowell, P. A. Rona, *Geophys. Res. Lett.* **29**, 1531 (2002).
12. G. Proskurowski, M. D. Lilley, D. S. Kelley, E. J. Olson, *Chem. Geol.* **229**, 331 (2006).
13. M. D. Lilley, D. A. Butterfield, J. E. Lupton, E. J. Olson, *Nature* **422**, 878 (2003).
14. W. E. Seyfried Jr., D. I. Foustoukos, Q. Fu, *Geochim. Cosmochim. Acta* **71**, 3872 (2007).
15. B. Martin, W. S. Fyfe, *Chem. Geol.* **6**, 185 (1970).
16. M. D. Lilley *et al.*, *Nature* **364**, 45 (1993).
17. J. A. Welhan, J. E. Lupton, *AAPG Bull.* **71**, 215 (1987).
18. M. J. Whiticar, *Chem. Geol.* **161**, 291 (1999).
19. B. Sherwood Lollar *et al.*, *Chem. Geol.* **226**, 328 (2006).
20. J. M. Hunt, *Petroleum Geochemistry and Geology* (Freeman, New York, 1996).
21. Y. A. Taran, G. A. Kliger, V. S. Sevastianov, *Geochim. Cosmochim. Acta* **71**, 4474 (2007).
22. B. Sherwood Lollar, T. D. Westgate, J. A. Ward, G. F. Slater, C. G. Lacrampe, *Nature* **416**, 522 (2002).
23. A. Schimmelmann, A. L. Sessions, M. Mastalerz, *Annu. Rev. Earth Planet. Sci.* **34**, 501 (2006).
24. W. F. Giggenbach, *Geochim. Cosmochim. Acta* **61**, 3763 (1997).
25. J. Horita, M. E. Berndt, *Science* **285**, 1055 (1999).
26. P. A. Rona, L. Widenfalk, K. Bostroem, *J. Geophys. Res.* **92**, 1417 (1987).
27. G. L. Früh-Green, J. A. D. Connolly, A. Plas, D. S. Kelley, B. Grobety, in *The Subseafloor Biosphere at Mid-Ocean Ridges*, W. S. D. Wilcock, E. F. Delong, D. S. Kelley, J. A. Baross, C. S. Cary, Eds. (American Geophysical Union, Washington, DC, 2004), vol. 144, pp. 119–136.
28. Low concentrations of CO_2 with large contributions from modern seawater bicarbonate prevent the determination of presumably near–radiocarbon dead end-member ^{14}C contents of Lost City fluids.
29. G. Proskurowski, M. D. Lilley, T. A. Brown, *Earth Planet. Sci. Lett.* **225**, 53 (2004).
30. J. C. Alt, D. A. H. Teagle, *Geochim. Cosmochim. Acta* **63**, 1527 (1999).
31. B. Marty, I. N. Tolstikhin, *Chem. Geol.* **145**, 233 (1998).
32. Seawater bicarbonate accounts for 55 to 98% of the measured CO_2 concentrations in LCHF fluids. The large extrapolations involved in the calculation of end-member CO_2 concentrations introduce large errors that propagate during the determination of end-member $\delta^{13}C$-CO_2 values by mass-isotope balance. However, measured $\delta^{13}C$-CO_2 values were −5 to −2‰, indicating mixing between seawater bicarbonate (−0.40‰) and an end-member more depleted in ^{13}C. The best-quality sample yielded an end-member $\delta^{13}C$-CO_2 value of −8.7 ± 1.5‰, and samples with greater uncertainties had lower values.
33. D. S. Kelley, G. L. Früh-Green, *Geochim. Cosmochim. Acta* **65**, 3325 (2001).
34. W. J. Brazelton, M. O. Schrenk, D. S. Kelley, J. A. Baross, *Appl. Environ. Microbiol.* **72**, 6257 (2006).
35. M. O. Schrenk, D. S. Kelley, S. A. Bolton, J. A. Baross, *Environ. Microbiol.* **6**, 1086 (2004).
36. D. S. Kelley, G. L. Früh-Green, *J. Geophys. Res.* **104**, 10439 (1999).
37. A. Delacour, P. Schaeffer, S. M. Bernasconi, G. L. Früh-Green, *Eos Trans. AGU* **87**, B31B (2006).
38. We thank the captains and crews of the R/V *Atlantis* and R/V *Ronald H. Brown*, and the crews of the DSV *Alvin* and the ROV *Hercules* for their indispensable expertise in deep-sea oceanography. We also thank R. Ballard and the Institute for Exploration for their efforts during the 2005 Lost City expedition, a proof-of-concept cruise directed by satellite-linked shore-based investigators. This work was supported in part by NSF grant OCE0137206, the NOAA Ocean Exploration Program, and a fellowship from the Woods Hole Oceanographic Institution Deep Ocean Exploration Institute. We thank D. Butterfield for providing Mg data. This manuscript benefited greatly from the comments by two anonymous reviewers.

Prioritizing Climate Change Adaptation Needs for Food Security in 2030

David B. Lobell,[1,2]* Marshall B. Burke,[1] Claudia Tebaldi,[3] Michael D. Mastrandrea,[4] Walter P. Falcon,[1] Rosamond L. Naylor[1]

Investments aimed at improving agricultural adaptation to climate change inevitably favor some crops and regions over others. An analysis of climate risks for crops in 12 food-insecure regions was conducted to identify adaptation priorities, based on statistical crop models and climate projections for 2030 from 20 general circulation models. Results indicate South Asia and Southern Africa as two regions that, without sufficient adaptation measures, will likely suffer negative impacts on several crops that are important to large food-insecure human populations. We also find that uncertainties vary widely by crop, and therefore priorities will depend on the risk attitudes of investment institutions.

Adaptation is a key factor that will shape the future severity of climate change impacts on food production (1). Although relatively inexpensive changes, such as shifting planting dates or switching to an existing crop variety, may moderate negative impacts, the biggest benefits will likely result from more costly measures including the development of new crop varieties and expansion of irrigation (2). These adaptations will require substantial investments by farmers, governments, scientists, and development organizations, all of whom face many other demands on their resources. Prioritization of investment needs, such as through the identification of "climate risk hot spots" (3), is therefore a critical issue but has received limited attention to date.

We consider three components to be essential to any prioritization approach: (i) selection of a time scale over which impacts are most relevant to investment decisions, (ii) a clear definition of criteria used for prioritization, and (iii) an ability to evaluate these criteria across a suite of crops and regions. Here, we focus on food security impacts by 2030: a time period most relevant to large agricultural investments, which typically take 15 to 30 years to realize full returns (4, 5).

We consider several different criteria for this time scale. First is the importance of the

[1]Food Security and Environment Program, Woods Institute for the Environment and the Freeman Spogli Institute for International Studies, Stanford University, Stanford, CA 94305, USA. [2]Lawrence Livermore National Laboratory (LLNL), Livermore, CA 94550, USA. [3]National Center for Atmospheric Research, Boulder, CO 80305, USA. [4]Woods Institute for the Environment, Stanford University, Stanford, CA 94305, USA.

*To whom correspondence should be addressed. E-mail: dlobell@stanford.edu

crop to a region's food-insecure human population [hunger importance (HI)]. Second is the median projected impact of climate change on a crop's production by 2030 (indicated by C50), assuming no adaptation. For this analysis, we generate multiple (i.e., 100) projections of impacts based on different models of climate change and crop response, in order to capture relevant uncertainties. The projections are then ranked, and the average of the 50th and 51st values are used as the median. A third criterion is the fifth percentile of projected impacts by 2030 (where C05 indicates the fifth value of the ranked projections), which we use to represent the lower tail or "worst case" among the projections. Finally, we consider the 95th percentile of projected impacts by 2030 (where C95 indicates the 95th value of the ranked projections), which we use to represent the upper tail or "best case" among the projections.

We first identified 12 major food-insecure regions, each of which (i) comprise groups of countries with broadly similar diets and agricultural production systems and (ii) contain a notable share of the world's malnourished individuals as estimated by the Food and Agriculture Organization (FAO) (Table 1; see fig. S1 for details on regions). For each region, we computed the HI value for each crop by multiplying the number of malnourished individuals by the crop's percent contribution to average per capita calorie consumption [see supporting online material (SOM) Text S1 and table S1]. A hunger importance ranking (HIR) was then generated by ranking the HI values for all crop-by-region combinations. Rice, maize, and wheat contribute roughly half of the calories currently consumed by the world's poor and only 31% of the calories consumed by those in sub-Saharan Africa, illustrating the importance of considering additional crops in food security assessments. The use of projected malnourished populations in 2030 rather than current population values had a very small influence on the rankings (table S2).

Several options exist for evaluating climate change impacts across a suite of crops and regions (SOM Text S2). We used data sets on historical crop harvests (6), monthly temperatures and precipitation, and maps of crop locations to develop statistical crop models for 94 crop-region combinations spanning the 12 study regions (see SOM Text S3; results summarized in Table 1). Of these combinations, 46% (43) exhibited a statistically significant model ($P < 0.05$), and 22% (21) had a model R^2 of at least 0.3. As seen in the examples for wheat in South and West Asia (fig. S3), in some cases the model's strength came primarily from a (typically negative) temperature effect on yield, whereas, in other cases, a (typically positive) rainfall effect provided most of the explanatory power.

The crop temperature sensitivities estimated by the statistical models were compared with corresponding values from previous studies that relied on established process-based models within the same regions (SOM Text S4). Our statistical estimates generally overlapped the lower end of the range of previous estimates, indicating that impacts estimated by the statistical models may be considered conservative but in reasonable agreement with estimates from process-based approaches.

To project climate changes for the crop regions, along with their uncertainties, we used output from 20 general circulation models (GCMs) that have contributed to the World Climate Research Programme's Coupled Model Intercomparison Project phase 3 (WCRP CMIP3) (7). Median projections of average temperature change from 1980–2000 to 2020–2040 were roughly 1.0°C in most regions, with few models projecting less than 0.5°C warming in any season and some models warming by as much as 2.0°C (Fig. 1A). In contrast to the unanimous warming, models were mixed in the direction of simulated precipitation change. All regions had at least one model with positive and one model with negative projected precipitation changes, with median projections ranging from about −10% to +5% (Fig. 1B). Some relevant tendencies of current GCMs, as noted in (8), are toward precipitation decreases during December to February (DJF) in South Asia and Central America, precipitation decreases in June to August (JJA) in Southern Africa, Central America, and Brazil, and precipitation increases in DJF in East Africa.

We estimated a probability distribution of production changes for 2030 (the average from

Table 1. Regions evaluated in this study and selected summary statistics. Countries within each region are indicated in the SOM.

| Region | Code | Malnourished | | Crops modeled | Crops with significant model* |
		Millions of people	World total (%)		
South Asia	SAS	262.6	30.1%	9	7
China	CHI	158.5	18.2%	7	2
Southeast Asia	SEA	109.7	12.6%	7	4
East Africa	EAF	79.0	9.1%	10	2
Central Africa	CAF	47.6	5.5%	8	0
Southern Africa	SAF	33.3	3.8%	8	6
West Africa	WAF	27.5	3.2%	8	2
Central America and Caribbean	CAC	25.4	2.9%	5	2
Sahel	SAH	24.9	2.9%	7	7
West Asia	WAS	21.9	2.5%	10	4
Andean region	AND	21.4	2.5%	9	3
Brazil	BRA	13.5	1.6%	6	4
Total	ALL	825.3	94.7%	94	43

*A model was judged significant if it explained more than 14% of variance in yield or production ($R^2 > 0.14$). This threshold was based on the 95th percentile of the R^2 statistic from a Monte Carlo experiment, which computed 1000 multiple regression models for a randomly generated 42-year time series with two random predictor variables.
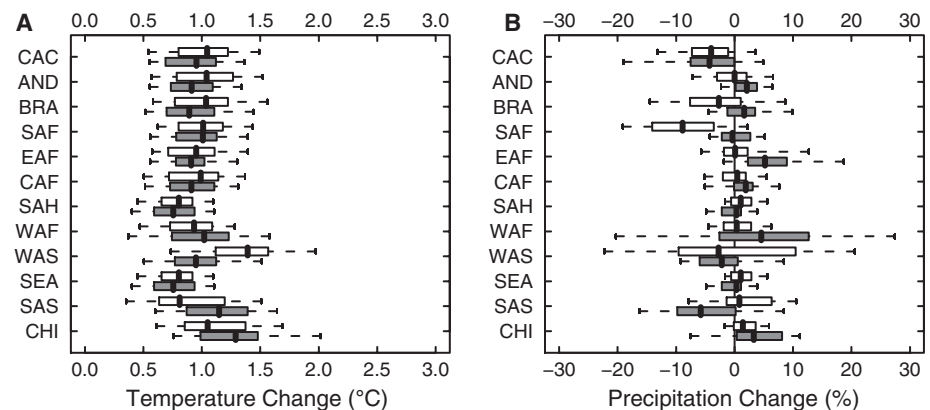
Fig. 1. Summary of projected (**A**) temperature (°C) and (**B**) precipitation (%) changes for 2030 (the averages from 2020 to 2039 relative to those from 1980 to 1999) based on output from 20 GCMs and three emission scenarios. Gray boxes show DJF averages and white boxes show JJA averages. Dashed lines extend from 5th to 95th percentile of projections, boxes extend from 25th to 75th percentile, and the middle vertical line within each box indicates the median projection.
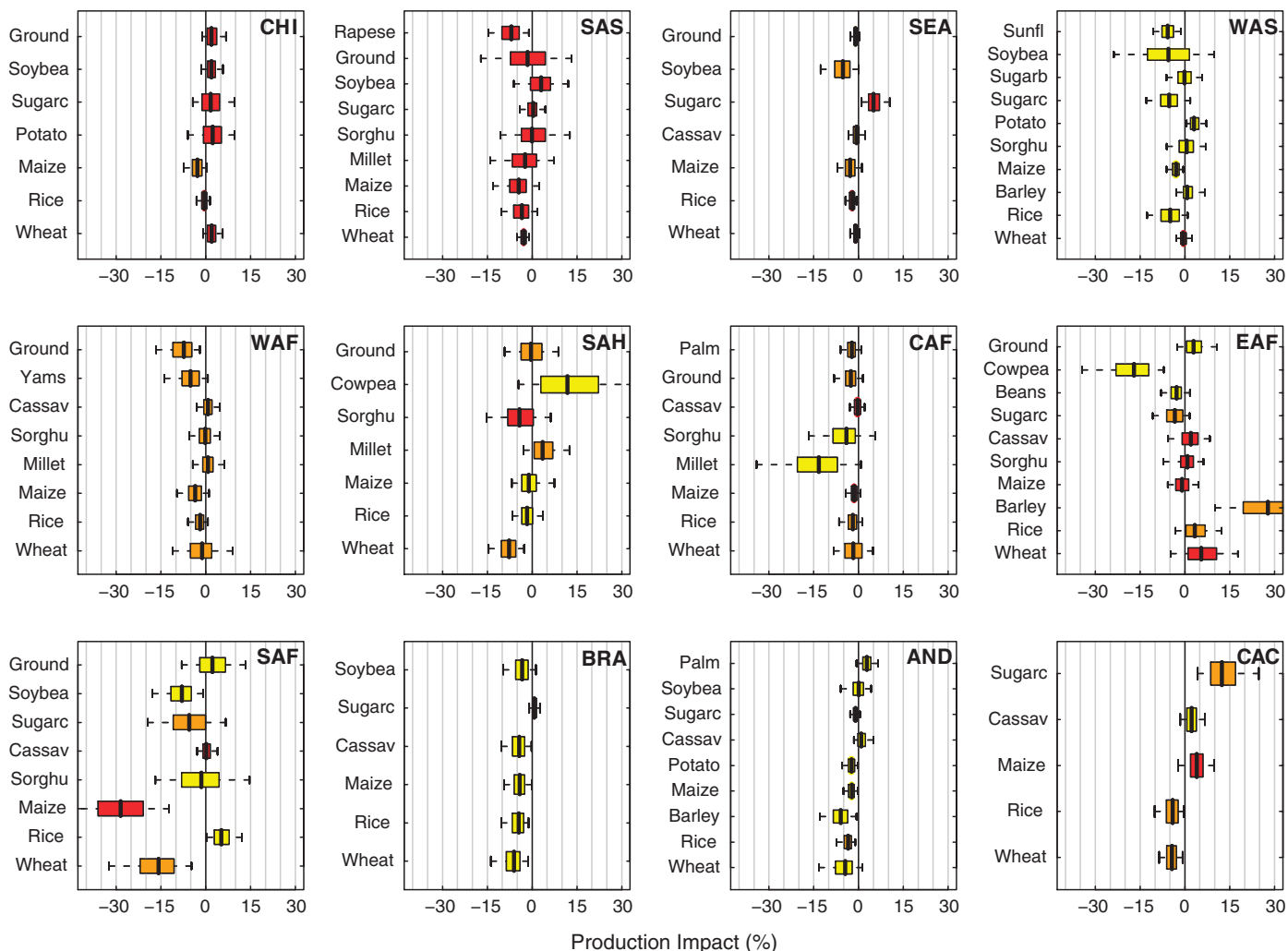
Fig. 2. Probabilistic projections of production impacts in 2030 from climate change (expressed as a percentage of 1998 to 2002 average yields). Red, orange, and yellow indicate a HIR of 1 to 30 (more important), 31 to 60 (important), and 61 to 94 (less important), respectively. Dashed lines extend from 5th to 95th percentile of projections, boxes extend from 25th to 75th percentile, and the middle vertical line within each box indicates the median projection. Region codes are defined in Table 1.

Table 2. Crop priority lists based on different criteria. C05 = 5th percentile of projected impacts (5th lowest out of 100 projections); C50 = 50th percentile (median); C95 = 95th percentile. Results are shown only for the HIR = 1 to 30 and HIR = 31 to 60 categories.

HIR value	Criterion	Crops
1 to 30	C05 < −10%	South Asia millet, groundnut, rapeseed; Sahel sorghum; Southern Africa maize
	C50 < −5%	South Asia rapeseed; Southern Africa maize
	C95 < 0%	South Asia wheat; Southeast Asia rice; Southern Africa maize
31 to 60	C05 < −10%	Southeast Asia soybean; West Asia rice; Western Africa wheat, yams, groundnut; Sahel wheat; East Africa sugarcane; Southern Africa wheat, sugarcane; Brazil wheat, rice; Andean Region wheat; Central America rice
	C50 < −5%	Southeast Asia soybean; West Asia rice; Western Africa yams, groundnut; Sahel wheat; Southern Africa wheat, sugarcane; Brazil wheat
	C95 < 0%	Western Africa groundnut; Sahel wheat; Southern Africa wheat; Brazil wheat, rice; Central America wheat, rice

2020 to 2039 relative to that from 1980 to 1999) for each crop using a Monte Carlo procedure that propagated both climate and crop uncertainties (*9*). To facilitate comparison between crops and regions, we expressed production changes for all crops as a percentage of average values for 1998 to 2002. The impact projections are summarized in Fig. 2.

For simplicity, we consider three general classes of projections. First, several projections (e.g., Southern Africa maize and wheat) are consistently negative, with an estimated 95% or greater chance that climate changes will harm crop production in the absence of adaptation (C95 < 0). These cases generally arise from a strong dependence of historical production variations on temperature, combined with projected warming large enough to overwhelm the uncertain impacts of precipitation changes.

Second, there are many cases with large uncertainties, with model impacts ranging from substantially negative to positive (e.g., South Asia groundnut, Southern Africa sorghum). These cases usually arise from a relatively strong dependence of historical production on rainfall, combined with large uncertainties in future precipitation changes. More precise projections of precipitation would therefore be particularly useful to reduce impact uncertainties in these situations. Large uncertainties also arise in some cases (e.g., cowpea in East Africa) from an estimated production response to historical temperature that is strongly negative but also highly uncertain.

Finally, there are many cases characterized by a narrow 90% confidence interval of impacts within ±5% of zero. In a few cases, such as wheat in West Asia, this reflects a strong effect of historical rainfall variations (fig. S1), combined with a relatively narrow range of rainfall projections during the growing season (Fig. 1; West Asia wheat is grown in DJF). In most cases, such as cassava in West Africa, the narrow confidence intervals result from a relatively weak relationship between historical production and growing-season climate. Therefore, we can only say that the likely impacts appear small, given the current data sets and models used to describe crop responses to climate. In cases with low model R^2, approaches other than the FAO-based regression models used here may be more appropriate.

Based on the above projections, we identified a small subset of crops that met different prioritization criteria (Table 2). First, crops were separated into groups of "more important" (HIR = 1 to 30), "important" (HIR = 31 to 60), and "less important" (HIR = 61 to 94). Within each category, we identified crops below three thresholds: the first corresponding to instances where at least 5% of the models predicted greater than 10% loss of production (C05 < –10%), the second to where at least half the models projected greater than a 5% production loss (C50 < –5%), and the third to where at least 95% of the models predicted some production loss (C95 < 0%).

Although several crops met more than one of these criteria, such as maize in Southern Africa and rapeseed in South Asia, the varying estimates of uncertainty for different crops, in general, resulted in noticeable differences when prioritizing crops on the basis of the three different thresholds (Table 2). For example, a relatively weak relationship was found between values at the two tails of the projection distributions—C05 and C95—across all crops (fig. S4). This result indicates a need to explicitly consider uncertainty and risk attitudes when setting priorities, which is an issue that has received limited attention (10).

Because attitudes toward risk differ, and given that impact projections for some crops are more uncertain than those for other crops, various institutions might derive different priorities from the results in Table 2. For example, one set of institutions might wish to focus on those cases where negative impacts are most likely to occur, in order to maximize the likelihood that investments will generate some benefits. By this criterion (C95 < 0%), South Asia wheat, Southeast Asia rice, and Southern Africa maize appear as the most important crops in need of adaptation investments.

Others might argue that adaptation activities that do not account for worst-case projections will be inadequate in the face of low-probability, high-consequence climate impacts: that is to say, investments should target those crops and re-

gions for which some models predict very negative outcomes. A different subset of crops is identified for this criterion (C05 < –10%), with several South Asian crops, Sahel sorghum, and (again) Southern Africa maize appearing as the most in need of attention.

Either of these risk attitudes could be applied with an explicit regional focus. For a sub-Saharan African institution interested in investing where negative impacts are most likely to occur [where median impact projections are substantially negative (C50 < –5%) or where most climate models agree that negative impacts are likely to occur (C95 < 0%)], priority investments would include Southern Africa maize, wheat, and sugarcane, Western Africa yams and groundnut, and Sahel wheat.

Despite the many assumptions and uncertainties associated with the crop and climate models used (SOM Text S5), the above analysis points to many cases where food security is clearly threatened by climate change in the relatively near-term. The importance of adaptation in South Asia and Southern Africa appears particularly robust, because crops in these regions appear for all criteria considered here (Table 2). The results also highlight several regions (e.g., Central Africa) where climate-yield relationships are poorly captured by current data sets, and therefore future work in this regard is needed to inform adaptation efforts.

Impacts will likely vary substantially within individual regions according to differences in biophysical resources, management, and other factors. The broad-scale analysis presented here was intended only to identify major areas of concern, and further studies at finer spatial scales are needed to resolve local hot spots within regions. Consideration of other social and technological aspects of vulnerability, such as the existing adaptive capacity in a region or the difficulty of making adaptations for specific cropping systems, should also be integrated into prioritization efforts. Although we do not attempt to identify the particular adaptation strategies that should be pursued, we note that, in some regions, switching from highly impacted to less impacted crops may be one viable adaptation option. In this case, the identification of less impacted crops is another valuable outcome of a comprehensive approach that simultaneously considers all crops relevant to the food-insecure.

References and Notes
1. W. Easterling et al., in Climate Change 2007: Impacts, Adaptation and Vulnerability. Contribution of Working Group II to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change (Cambridge Univ. Press, Cambridge, 2007), pp. 273–313.
2. C. Rosenzweig, M. L. Parry, Nature 367, 133 (1994).
3. I. Burton, M. van Aalst, "Look before you leap: A risk management approach for incorporating climate change adaptation in World Bank operations" (World Bank, Washington, DC, 2004).
4. J. M. Alston, C. Chan-Kang, M. C. Marra, P. G. Pardey, T. J. Wyatt, "A meta-analysis of rates of return to agricultural R&D: Ex pede Herculem?" (International Food Policy Research Institute, Washington, DC, 2000).
5. J. Reilly, D. Schimmelpfennig, Clim. Change 45, 253 (2000).
6. We used FAO data on national crop production and area, which include quantities consumed or used by the producers in addition to those sold on the market.
7. Model simulations under three SRES (Special Report on Emissions Scenarios) emission scenarios corresponding to relatively low (B1), medium (A1b), and high (A2) emission trajectories were used. Although the mean projections for the emission scenarios exhibit very small differences out to 2030, the use of three scenarios provided a larger sample of simulations with which to assess climate uncertainty. For all simulations, average monthly output for 1980–1999 was subtracted from that of 2020–2039 to compute monthly changes in temperature and precipitation.
8. J. H. Christensen et al., in Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change, S. Solomon et al., Eds. (Cambridge Univ. Press, Cambridge, 2007), pp. 847–940.
9. Namely, the crop regression model was fit with a bootstrap sample from the historical data, and the coefficients from the regression model were then multiplied by projected changes in average temperature and precipitation, which were randomly selected from the CMIP3 database. This process was repeated 100 times for each crop. Bootstrap resampling is a common approach to estimate uncertainty in regression coefficients, although this addresses only the component of model uncertainty that arises from a finite historical sample and not the potential uncertainty from structural errors in the model. Similarly, the representation of climate uncertainty by equally weighting all available GCMs is a common approach but could potentially be improved, such as by weighting models according to their agreement with the model consensus and with historical observations. Nonetheless, the resulting probability distributions incorporate reasonable measures of both climate and crop uncertainty, and thus should fairly reflect both the absolute and relative level of uncertainties between crops.
10. B. Smit et al., in Climate Change 2001: Impacts, Adaptation and Vulnerability. Contribution of Working Group II to the Third Assessment Report of the Intergovernmental Panel on Climate Change (Cambridge Univ. Press, Cambridge, 2001), pp. 877–912.
11. We thank D. Battisti, C. Field, and three anonymous reviewers for helpful comments. D.B.L. was supported by a Lawrence Fellowship from LLNL. Part of this work was performed under the auspices of the U.S. Department of Energy (DOE) by LLNL under contract DE-AC52-07NA27344. We acknowledge the modeling groups, the Program for Climate Model Diagnosis and Intercomparison, and the WCRP's Working Group on Coupled Modelling for their roles in making available the WCRP CMIP3 multimodel data set. Support of this data set is provided by the Office of Science, DOE.

Oocyte-Specific Deletion of *Pten* Causes Premature Activation of the Primordial Follicle Pool

Pradeep Reddy,[1] Lian Liu,[1,2]* Deepak Adhikari,[1]* Krishna Jagarlamudi,[1]* Singareddy Rajareddy,[1]* Yan Shen,[1] Chun Du,[1] Wenli Tang,[1] Tuula Hämäläinen,[3] Stanford L. Peng,[4] Zi-Jian Lan,[5] Austin J. Cooney,[6] Ilpo Huhtaniemi,[3,7] Kui Liu[1]†

In the mammalian ovary, progressive activation of primordial follicles from the dormant pool serves as the source of fertilizable ova. Menopause, or the end of female reproductive life, occurs when the primordial follicle pool is exhausted. However, the molecular mechanisms underlying follicle activation are poorly understood. We provide genetic evidence that in mice lacking PTEN (phosphatase and tensin homolog deleted on chromosome 10) in oocytes, a major negative regulator of phosphatidylinositol 3-kinase (PI3K), the entire primordial follicle pool becomes activated. Subsequently, all primordial follicles become depleted in early adulthood, causing premature ovarian failure (POF). Our results show that the mammalian oocyte serves as the headquarters of programming of follicle activation and that the oocyte PTEN-PI3K pathway governs follicle activation through control of initiation of oocyte growth.

Ovarian follicles are the basic units of a mammalian ovary, and each follicle contains an oocyte that is surrounded by somatic (granulosa) cells (*1*). Whereas the majority of follicles remain as dormant primordial follicles containing immature oocytes surrounded by a few flattened somatic cells for use in later reproductive life, a limited number of primordial follicles are recruited from the resting follicle reservoir into the growing follicle pool. The activation of primordial follicles is a progressive and highly regulated process. It ceases and female reproductive life ends when the follicle pool is exhausted at menopause (*1, 2*). The duration of fertility of a female is determined by the initial size of her primordial follicle pool and by the rate of its activation and depletion (*3*).

A landmark of follicular activation is the initiation of rapid oocyte growth within primordial follicles (*4, 5*). Activation is irreversible, such that activated follicles that are not selected for further development will undergo atresia. Although the initial follicular activation is believed to be independent of gonadotropin action, the molecular mechanisms are poorly defined (*1, 2*). Our recent studies (*5, 6*) suggest that the phosphatidylinositol 3-kinase (PI3K) pathway of oocytes may play a role in regulating the activation

of primordial follicles; however, further functional evidence is required to support this notion.

PI3Ks are lipid kinases that phosphorylate the 3′-OH group on the inositol ring of inositol phospholipids. PTEN (phosphatase and tensin homolog deleted on chromosome 10), a lipid phosphatase, reverses this process and thus functions as a major negative regulator of PI3K action (*7*). The PI3K pathway is a fundamental signaling pathway for the regulation of cell proliferation, survival, migration, and metabolism (*7*). To study the functional roles of the oocyte PI3K pathway in mammalian follicular activation, we deleted the *Pten* gene from mouse oocytes by crossing $Pten^{loxP/loxP}$ mice (*8*) with transgenic mice expressing *growth differentiation factor 9* (*Gdf-9*) promoter-mediated Cre recombinase (referred to as *GCre* mice); this recombinase is active specifically in oocytes (*9*) (fig. S1). We found that during a testing period from 6 to 34 weeks of age, the $Pten^{loxP/loxP};GCre^+$ females produced a maximum of one normal-sized litter but became infertile in early adulthood (i.e., after 12 to 13 weeks of age) (Fig. 1).

To study how the loss of *Pten* from oocytes impedes mouse fertility, we compared the first wave of postnatal follicular development in $Pten^{loxP/loxP};GCre^+$ and $Pten^{loxP/loxP}$ mice. We found no apparent morphological difference in postnatal day 5 (PD5) ovaries of $Pten^{loxP/loxP};GCre^+$ and $Pten^{loxP/loxP}$ mice. The ovaries of both genotypes had mostly primordial follicles containing small oocytes surrounded by flattened pregranulosa cells (Fig. 2, A to C, arrows) and some activated follicles containing enlarged oocytes (Fig. 2, A to C, arrowheads), with comparable numbers (fig. S3A). At PD8, however, the $Pten^{loxP/loxP};GCre^+$ ovaries appeared larger (fig. S2, A and B) with more activated follicles (fig. S3B), including transient follicles containing enlarged oocytes surrounded by flattened pregranulosa cells (fig. S2C, yellow arrows), primary (type 3b) follicles with enlarged oocytes sur-

rounded by one layer of cuboidal granulosa cells, and some secondary (type 4) follicles with two layers of granulosa cells (fig. S2, B and C, black arrows), whereas the $Pten^{loxP/loxP}$ ovaries had mostly primordial (fig. S2A, red arrows) and type 3b follicles (fig. S2A, arrowheads). Accordingly, the percentage of primordial follicles in $Pten^{loxP/loxP};GCre^+$ ovaries at PD8 (49.6%) was significantly lower than that in $Pten^{loxP/loxP}$ ovaries (83.6%) (fig. S3B). By PD23 and 35, the $Pten^{loxP/loxP};GCre^+$ ovaries (Fig. 2E and fig. S2E) remained larger than the $Pten^{loxP/loxP}$ ovaries (Fig. 2D and fig. S2D) and contained significantly more activated follicles (Fig. 2F, fig. S2F, and fig. S3, C and D). By PD23, virtually no primordial follicles could be identified in mutant ovaries (figs. S2E and S3C), whereas 69.2% of the follicles in control ovaries were still at the primordial stage (fig. S2D, inset, and fig. S3C). Therefore, the entire pool of primordial follicles had been activated in $Pten^{loxP/loxP};GCre^+$ ovaries. This also resulted in reduced follicle death and clearance before and around the time of sexual maturity in the mutant mice (*10*) (fig. S3F). The growth dynamics of the activated transient follicles seemed different: Some follicles appeared to remain at the transient stage (Fig. 2F and fig. S2F, yellow arrows; fig. S3, C to E), whereas others appeared to develop further (fig. S3, C to E).

At 7 weeks, $Pten^{loxP/loxP};GCre^+$ ovaries still appeared larger than $Pten^{loxP/loxP}$ ovaries (fig. S2, G and H), with significantly higher numbers of transient and preantral type 5 follicles (with three or more layers of granulosa cells) accumulated (fig. S3E). Corpora lutea (CL), which are temporary endocrine structures developed from the remnants of ovulated follicles, were present (fig. S2H, arrows), indicating that ovulation had taken place. Furthermore, in ovaries of 12-week-old $Pten^{loxP/loxP};GCre^+$ mice, follicular structure was largely deformed (fig. S2K) and CL were undergoing degeneration (luteolysis) (fig. S2K, arrows). In ovaries of both 7- and 12-week-old $Pten^{loxP/loxP};GCre^+$ mice, many transient follicles containing degraded oocytes were observed (fig. S2, I and L, arrows), which suggests that some of the prematurely activated follicles undergo atre-

[1]Department of Medical Biochemistry and Biophysics, Umeå University, SE-901 87 Umeå, Sweden. [2]Department of Chemotherapy, Cancer Center, Qilu Hospital, Shandong University, Jinan 250012, China. [3]Department of Physiology, Institute of Biomedicine, University of Turku, 20520 Turku, Finland. [4]Clinical Research and Exploratory Development, Roche Palo Alto, Palo Alto, CA 94304, USA. [5]Birth Defects Center, Department of Molecular, Cellular and Craniofacial Biology, University of Louisville Health Sciences Center, Louisville, KY 40202, USA. [6]Department of Molecular and Cellular Biology, Baylor College of Medicine, Houston, TX 77030, USA. [7]Department of Reproductive Biology, Imperial College London, Hammersmith Campus, London W12 0NN, UK.

*These authors contributed equally to this work.
†To whom correspondence should be addressed. E-mail: kui.liu@medchem.umu.se
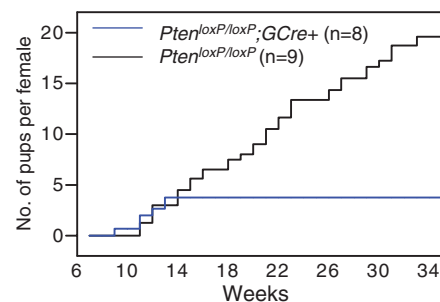
Fig. 1. Comparison of the cumulative number of pups per $Pten^{loxP/loxP};GCre^+$ female (blue line) and per $Pten^{loxP/loxP}$ female (black line); *n* = 9 for $Pten^{loxP/loxP}$ mice, *n* = 8 for $Pten^{loxP/loxP};GCre^+$ mice. All $Pten^{loxP/loxP};GCre^+$ females became infertile in young adulthood (after week 12 to 13).

sia directly from the transient stages. Mice at this or older ages showed completely irregular estrous cycles. By the age of 16 weeks, $Pten^{loxP/loxP};GCre^+$ ovaries no longer displayed normal ovarian morphology (Fig. 2, H and I). The ovaries were smaller, rounded, and yellow (Fig. 2H, inset) in comparison to the pink-colored $Pten^{loxP/loxP}$ ovaries, which contained CL (Fig. 2G, arrows) and follicles (Fig. 2G, arrowheads).

Moreover, in sera of 12- to 20-week-old $Pten^{loxP/loxP};GCre^+$ mice, elevated levels of follicle-stimulating hormone (FSH) (Fig. 3A) and luteinizing hormone (LH) (Fig. 3B) were observed relative to $Pten^{loxP/loxP}$ control mice. Thus, activation of the pool of primordial follicles leads to follicle depletion; this causes premature ovarian failure (POF) in $Pten^{loxP/loxP};GCre^+$ mice. The phenotype observed in these mice resembles that of human POF (11).

To elucidate the molecular mechanisms underlying the accelerated oocyte enlargement in $Pten^{loxP/loxP};GCre^+$ ovaries, we studied Akt signaling in oocytes isolated from ovaries of PD12 to 14 $Pten^{loxP/loxP};GCre^+$ and $Pten^{loxP/loxP}$ mice. We found that the level of phospho-Akt (p-Akt, Ser^{473}) was elevated in $Pten^{loxP/loxP};GCre^+$ oocytes that were cultured in vitro and starved of serum (Fig. 4A). In addition, Kit ligand (KL), which can activate the PI3K pathway in growing oocytes through its oocyte surface receptor Kit (6), activated Akt to a greater extent in $Pten^{loxP/loxP};GCre^+$ oocytes than in $Pten^{loxP/loxP}$ oocytes (Fig. 4B). Thus, the loss of Pten in oocytes leads to enhanced oocyte PI3K-Akt signaling.

To investigate the cause of accelerated oocyte growth in $Pten^{loxP/loxP};GCre^+$ ovaries, we studied whether the enhanced PI3K-Akt signaling led to increased activation of ribosomal protein S6 (rpS6). At PD5—a developmental stage with no apparent morphological differences between $Pten^{loxP/loxP};GCre^+$ and $Pten^{loxP/loxP}$ ovaries (Fig. 2, A to C)—the activation of Akt had already been elevated in $Pten^{loxP/loxP};GCre^+$ oocytes (Fig. 4C, PD5, p-Akt). This result correlated with enhanced expression (Fig. 4C, PD5, rpS6) and phosphorylation (indicating activation) of rpS6 (Fig. 4C, PD5, p-rpS6, $Ser^{235/6}$). Such a result suggests that enhanced protein translation had already started when the Gdf-9–Cre-mediated Pten deletion in oocytes had just taken place (9). Similarly, in oocytes isolated from $Pten^{loxP/loxP};GCre^+$ ovaries at PD12 to 14, enhanced PI3K-Akt signaling led to elevation of both expression and phosphorylation of rpS6 (Fig. 4C, PD12 to 14). However, activation of the mammalian target of rapamycin (mTOR)–p70 S6 kinase (S6K) cascade was not increased by loss of Pten, as the levels of phospho-mTOR (p-mTOR, Ser^{2448}), phospho-tuberin/TSC2 (p-TSC2, Thr^{1462}), and phospho-S6K (p-S6K, Thr^{389}) at PD5 and PD12 to 14 remained similar in $Pten^{loxP/loxP};GCre^+$ and $Pten^{loxP/loxP}$ oocytes (Fig. 4C). Thus, the enhanced activation of rpS6 was caused by elevated rpS6 expression per se (Fig. 4C). However, the phosphorylation of rpS6

and S6K in mutant oocytes was sensitive to the PI3K-specific inhibitor LY294002 and the mTOR-specific inhibitor rapamycin (fig. S4A), indicating that activation of rpS6 in $Pten^{loxP/loxP};GCre^+$ oocytes is dependent on the activities of PI3K and mTOR.

It has been hypothesized that unknown intra-ovarian factors stimulate some primordial follicles to initiate growth while the rest of the follicles remain quiescent. On the other hand, it has also been suggested that follicular activation is triggered by the release of inhibitory mechanisms that maintain the primordial follicles in their

resting state (1). Our data show that the oocyte governs follicular activation and that oocyte PTEN functions as a suppressor of this process. The intra-oocyte PTEN-PI3K signaling cascade appears to play a role in the initiation of oocyte growth. We propose that activation of the PI3K pathway in each individual oocyte may be essential in determining the fate of the primordial follicle—whether it remains dormant, whether it becomes activated at a certain time, or whether it undergoes atresia directly from the primordial stage.

Furthermore, we report the distinctive ovarian phenotype of POF in mice with oocyte-specific

$Pten^{loxP/loxP}$ $Pten^{loxP/loxP};GCre+$

Fig. 2. Activation of the primordial follicle pool followed by premature ovarian failure (POF) in $Pten^{loxP/loxP};GCre^+$ mice. (**A** to **F**) Overactivation of primordial follicles in $Pten^{loxP/loxP};GCre^+$ mice. Ovaries from PD5 and PD35 $Pten^{loxP/loxP};GCre^+$ mice and their $Pten^{loxP/loxP}$ littermates were embedded in paraffin, and sections of 8-μm thickness were prepared and stained with hematoxylin. (**G** to **I**) POF caused by depletion of all follicles in $Pten^{loxP/loxP};GCre^+$ mice. Ovaries from 16-week-old $Pten^{loxP/loxP};GCre^+$ mice and $Pten^{loxP/loxP}$ littermates were used. The experiments were repeated more than three times, and for each time and each age, ovaries from one mouse of each genotype were used.

Fig. 3. Elevated levels of FSH (**A**) and LH (**B**) in adult $Pten^{loxP/loxP};GCre^+$ mice (mutant) and $Pten^{loxP/loxP}$ mice (control); n = 14 for $Pten^{loxP/loxP};GCre^+$ mice, n = 13 for $Pten^{loxP/loxP}$ mice. Female mice of both genotypes (12 to 20 weeks old) were killed and sera were collected for measurement of FSH and LH levels (10). P values are shown in each panel.

Fig. 4. Enhanced Akt signaling in $Pten^{loxP/loxP};GCre^+$ oocytes leads to elevated expression and activation of rpS6. Oocytes were isolated from ovaries of $Pten^{loxP/loxP};GCre^+$ and $Pten^{loxP/loxP}$ mice at PD5 and PD12 to 14, and Western blots were performed (10). (**A**) Levels of p-Akt (Ser[473]) and total Akt in PD12 to 14 $Pten^{loxP/loxP};GCre^+$ and $Pten^{loxP/loxP}$ oocytes. (**B**) Activation of Akt (p-Akt, Ser[473]) by KL treatment (100 ng/ml, 2 min) in PD12 to 14 $Pten^{loxP/loxP};GCre^+$ and $Pten^{loxP/loxP}$ oocytes. Levels of Akt were used as internal controls. (**C**) Signaling studies in $Pten^{loxP/loxP};GCre^+$ and $Pten^{loxP/loxP}$ oocytes at PD5 and PD12 to 14, showing levels of p-Akt (Ser[473]), rpS6, p-rpS6 (Ser[235/6]), p-mTOR (Ser[2448]), p-TSC2 (Thr[1462]), and p-S6K (Thr[389]). Levels of total Akt, mTOR, TSC2, S6K, and β-actin were used as internal controls. All experiment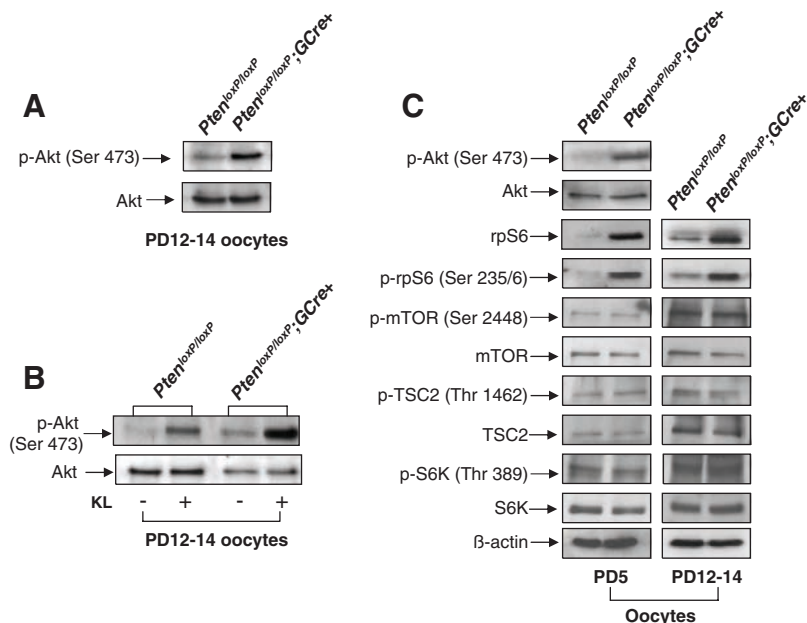s were repeated at least three times. For isolation of PD5 oocytes for Western blot, 10 to 15 $Pten^{loxP/loxP};GCre^+$ or $Pten^{loxP/loxP}$ mice were used for each lane. For isolation of PD12 to 14 oocytes, 3 to 5 $Pten^{loxP/loxP};GCre^+$ mice or 6 to 10 $Pten^{loxP/loxP}$ mice were used per lane. In each lane, 30 to 40 μg of protein sample was loaded.

ablation of *Pten*, which is caused by excessive activation and depletion of primordial follicles. Thus, our findings may have broad physiological and clinical implications, contributing to in-depth understanding of both normal ovarian physiology and the development of ovarian diseases. In humans, POF is defined as a primary ovarian defect characterized by absent menarche (primary amenorrhea) or by premature depletion of ovarian follicles or arrested folliculogenesis be-

fore the age of 40 years (secondary amenorrhea), with an estimated incidence of 1% (*11*). We hypothesize that genetic variations leading to overactivation and depletion of follicles may be among the possible causes of POF in humans. On the other hand, the retardation of follicle activation and/or excessive primordial follicle atresia, both of which may be caused by underactivation of the PI3K pathway in oocytes, can also lead to POF, albeit from opposite directions. Recognition of the importance of the PTEN-PI3K signaling network in oocytes opens up new prospects for our understanding of the physiological and pathological processes of the mammalian ovary.

References and Notes

1. E. A. McGee, A. J. Hsueh, *Endocr. Rev.* **21**, 200 (2000).
2. A. N. Hirshfield, *Int. Rev. Cytol.* **124**, 43 (1991).
3. G. F. Erickson, *J. Soc. Gynecol. Investig.* **8**, S13 (2001).
4. P. M. Wassarman, D. F. Albertini, in *The Physiology of Reproduction*, E. Knobil, J. D. Neill, Eds. (Raven, New York, 1994), vol. 1, pp. 79–122.
5. K. Liu *et al.*, *Dev. Biol.* **299**, 1 (2006).
6. P. Reddy *et al.*, *Dev. Biol.* **281**, 160 (2005).
7. L. C. Cantley, *Science* **296**, 1655 (2002).
8. M. Groszer *et al.*, *Science* **294**, 2186 (2001); published online 1 November 2001 (10.1126/science.1065518).
9. Z. J. Lan, X. Xu, A. J. Cooney, *Biol. Reprod.* **71**, 1469 (2004).
10. See supporting material on *Science* Online.
11. P. Beck-Peccoz, L. Persani, *Orphanet. J. Rare. Dis.* **1**, 9 (2006).
12. Supported by the J. C. Kempe and Seth M. Kempe Memorial Foundation, the Swedish Research Council, the Swedish Cancer Foundation, Lion's Cancer Research Foundation at Umeå University, and the Novo Nordisk Foundation (K.L.).

The Maternal Nucleolus Is Essential for Early Embryonic Development in Mammals

Sugako Ogushi,[1,2,3]* Chiara Palmieri,[5] Helena Fulka,[3,4] Mitinori Saitou,[2] Takashi Miyano,[1] Josef Fulka Jr.[3]

With fertilization, the paternal and maternal contributions to the zygote are not equal. The oocyte and spermatozoon are equipped with complementary arsenals of cellular structures and molecules necessary for the creation of a developmentally competent embryo. We show that the nucleolus is exclusively of maternal origin. The maternal nucleolus is not necessary for oocyte maturation; however, it is necessary for the formation of pronuclear nucleoli after fertilization or parthenogenetic activation and is essential for further embryonic development. In addition, the nucleolus in the embryo produced by somatic cell nuclear transfer originates from the oocyte, demonstrating that the maternal nucleolus supports successful embryonic development.

To create a totipotent zygote, the oocyte and spermatozoon combine not only their nuclear DNA but also certain RNAs, proteins, and organelles. Some zygotic material is strictly of maternal or paternal origin. For example, mitochondria originate exclusively from the oocyte (*1*), whereas the centriole in most mammals (excluding rodents) comes from the sper-

matozoon (*2, 3*). Another zygotic organelle that seems to be supplied by the oocyte is the nucleolus. The nucleoli in fully grown oocytes are compact and transcriptionally inactive (*4–7*). It is not known how or to what extent this nucleolar material contributes to the construction of zygotes and early embryos. Our results demonstrate that the nucleolus in the zygote and early embryo is exclusively maternally inherited and originates from the material that is present in the oocyte germinal vesicle. Moreover, using nuclear transfer experiments, we demonstrate that nucleoli originating from a somatic cell or even from an embryonic stem (ES) cell cannot substitute for the original oocyte nucleolar material.

The scheme for our analysis is shown in fig. S1. Nucleoli were microsurgically removed from fully grown oocytes before gonadotropic stimulations in pigs and mice (Fig. 1, A and B, fig. S2, and movies S1 and S2) (*8*). The oocytes from which we aspirated a small amount of nucleoplasm served as controls [fig. S1, sham-operated (Sham)]. The success of enucleolation was checked immunocytochemically (Fig. 1C and fig.

S3A). In sham-operated oocytes, the immuno-fluorescence against H3K9triMet (trimethylated histone H3 at Lys[9]), C23 (nucleolin), and NPM2 (nucleoplasmin 2), as well as the staining against chromatin by DAPI (4´,6´-diamidino-2-phenylindole) and Hoechst (bisbenzimide H33342), demonstrated a single nucleolus surrounded by a heterochromatin rim (Fig. 1C, Chromatin, and fig. S3A, Hoechst). With this approach, no nucleoli were detected in enucleolated oocytes. After the enucleolation, the germinal vesicle envelope became sealed, as confirmed by labeling against lamin A/C (nuclear lamin A and lamin C) (Fig. 1C and fig. S3B, Lamin A/C), and the chromatin remained inside the germinal vesicle and was not detectable around isolated nucleoli (Fig. 1C, Chromatin, and fig. S3A, Hoechst). Maturation of enucleolated oocytes appeared normal, with regular spindle formation and subsequent polar body extrusion, and progressed in a similar time course as in controls to metaphase II at 52% (93/180) in the pig (fig. S3, D and F) and 90% (533/592) in the mouse (fig. S3B). Karyotyping of mouse oocytes at metaphase II revealed a normal haploid set of chromosomes in both sham-operated and enucleolated oocytes (n = 20 chromosomes, fig. S3C). Analysis of the kinetics of activities of meiotic cell cycle molecules cyclin-dependent kinase 1 (CDK1) and mitogen-activated protein (MAP) kinase demonstrates that these kinase activities fluctuated similarly in the enucleolated and control pig oocytes (fig. S3E). These results suggest that oocyte nucleoli do not contain the essential factor(s) for oocyte maturation.

In both species, enucleolated and mature oocytes were efficiently activated partheno-genetically (pig: 100%, 212/212; mouse: 75%, 157/208) or by fertilization (pig: 57%, 93/164; mouse: 69%, 129/186). Whereas round prominent nucleoli were detectable in germinal vesicles and in both the male and female pronuclei of control zygotes, they were not visible in germinal vesicles or pronuclei originating from enucleolated oocytes (Fig. 1, D to F). After activation, CDK1 and MAP kinase in enucleolated oocytes were inactivated in a similar time course as in controls (fig. S3E). At the pronucleus or four- to eight-cell stage, the proteins and DNA were synthesized at similar levels in both the experimental and control groups (fig. S4).

Next, we tested the recovery of nucleolus formation in the pronuclei by reinjecting the oocyte nucleoli into the enucleolated oocytes

(fig. S1, Nucleolus reinjected). Isolated pig oocyte nucleoli were kept for 26 hours in the empty zonae pellucidae prepared by sucking out all the contents of the ooplasm by a micromanipulator (Fig. 1B), and mouse oocyte nucleoli were kept in the perivitelline space under the zona pellucida for 14 hours (fig. S3A). The nucleoli were then reinjected into enucleolated mature oocytes. Reinjected nucleoli disappeared in the oocyte cytoplasm (fig. S5). After activation, the nucleolus-reinjected oocytes formed pronuclei that contained prominent nucleoli (pig: 53%, 61/115; mouse: 69%, 77/112), whereas no nucleoli were visible in embryos from enucleolated oocytes (Fig. 2 and fig. S6). In addition, after in vitro fertilization, zygotes from the nucleolus-reinjected enucleolated oocytes contained nucleoli in their pronuclei (fig. S7). Next, we tested whether somatic cell nucleoli can substitute for the original oocyte material by injecting a somatic cell

nucleus via the somatic cell nuclear transfer (SCNT) method. After enucleolation and maturation of pig and mouse oocytes, nuclei of cumulus cells were injected into cytoplasts of enucleated oocytes at metaphase II (fig. S1, Enucleolated-SCNT). Cytoplasts from non-enucleolated mature oocytes served as controls (fig. S1, Sham-SCNT). The exposure of somatic nuclei to the oocyte cytoplasm resulted in nuclear envelope breakdown and premature chromosome condensation (pig: 59%, 73/123; mouse: 73%, 38/52). These reconstructed oocytes were then activated parthenogenetically. No nucleoli were seen in newly formed pseudo-pronuclei in the enucleolated group, whereas controls always contained a single nucleolus or multiple nucleoli (Fig. 2). Similarly, no nucleoli were visible in pseudo-pronuclei after injection of nuclei of pluripotent ES cells into enucleolated oocytes (Fig. 2B, Enucleolated-ESCNT). This indicates that somatic and ES
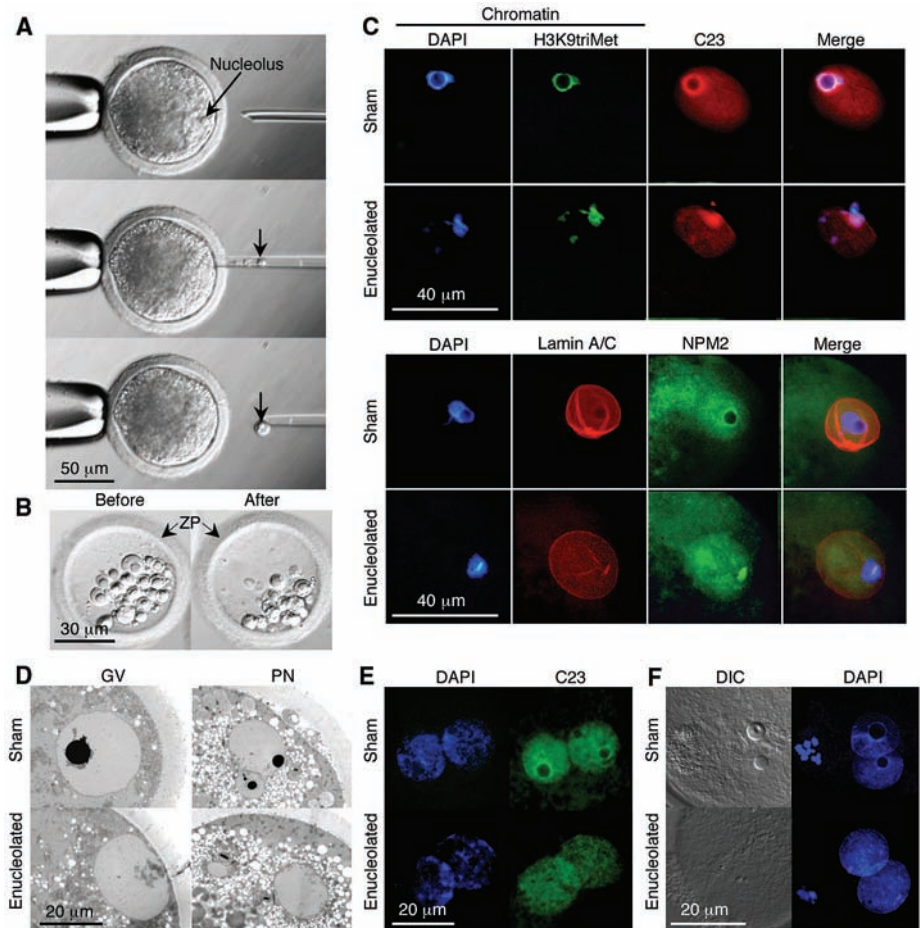
Fig. 1. Enucleolation of immature oocytes (A to C) and the absence of nucleoli in pronuclei of zygotes originating from enucleolated oocytes (D to F). (**A**) The nucleolus in centrifuged pig oocytes at the germinal vesicle (GV) stage was aspirated and isolated. (**B**) Isolated nucleoli before (left) and 26-hours after (right) culture without disassembling. Arrows indicate the empty zonae pellucidae (ZP). (**C**) Immunofluorescence labeling was performed with the indicated antibodies in green (H3K9triMet and NPM2) and in red (C23 and Lamin A/C). The absence of a nucleolus is evident in enucleolated oocytes. (**D** to **F**) No nucleolus formation in pronuclei (PN) is observed in light-microscopic observations (D) and immunofluorescence labeling against C23 (green) (E) of pig zygotes, as well as in differential interference contrast (DIC) images of mouse zygotes after in vitro fertilization (F). DAPI staining marks chromatin in blue.

[1]Graduate School of Science and Technology, Kobe University, Kobe 657-8501, Japan. [2]Laboratory for Mammalian Germ Cell Biology, RIKEN Kobe Institute, Kobe 650-0047, Japan. [3]Institute of Animal Science, Prague 104 01, Czech Republic. [4]Institute of Molecular Genetics, Academy of Sciences of the Czech Republic, Prague 142 20, Czech Republic. [5]Faculty of Veterinary Medicine, University of Teramo, Teramo 64100, Italy.

*To whom correspondence should be addressed. E-mail: ogushi@cdb.riken.jp

cell nuclei (including nucleolus) cannot substitute for the original oocyte nucleolar material.

The development of pig embryos originating from enucleolated oocytes arrested after a few cleavages (Table 1 and Fig. 3A). In contrast, embryos from sham-operated oocytes developed to blastocysts with detectable nucleoli (fig. S8, Pig). The cleavage rate was essentially the same at 48 hours after activation. However, no nucleoli were detected in the nuclei of two-cell embryos originating from enucleolated oocytes (Fig. 3B, Pig). At this stage, these embryos without nucleoli were still viable, having the ability to synthesize proteins at similar levels when compared to control embryos originating from sham-operated oocytes (fig. S4, A and B). We observed the same developmental defects in the embryos from enucleolated oocytes after SCNT (Table 1 and Fig. 3B, Pig), and intracytoplasmic sperm injection (table S1). Unexpectedly, pig embryos that were enucleolated and then reinjected with oocyte nucleoli cleaved regularly with visible nucleoli in blastomere nuclei (Fig. 3B, Pig) and developed to blastocysts (Table 1, table S1, Fig. 3A, and fig. S8, Nucleolus reinjected). The developmental ability of embryos, including SCNT embryos, thus depends on the presence of the oocyte nucleolar material that was either released into the oocyte cytoplasm at germinal vesicle breakdown or reinjected into the cytoplasm of previously enucleolated oocytes. In the mouse (table S2), embryos originating from sham-operated oocytes developed to the blastocyst stage, and the blastomere nuclei contained visible nucleoli (fig. S8, Mouse). No nucleoli were detected in embryos originating from enucleolated oocytes (Fig. 3B, Mouse), and their development was arrested between the two- and four-cell stages. After nucleolus reinjection into enucleolated and mature mouse oocytes, the nucleoli were detected in embryo nuclei, and their developmental potential was restored (table S2). Furthermore, when we transferred two-cell embryos that were derived from enucleolated oocytes and from nucleolus-reinjected oocytes to recipients' uteri, we obtained live-born pups at a ratio comparable to that from control oocytes (table S3). Thus, the inability of enucleolated oocytes to develop does not reflect the mechanical damage caused by micromanipulation but is a direct consequence of the absence of nucleolar material.

Cytoplasts of oocytes at metaphase II and mitotic zygotes support development of SCNT embryos (9). In contrast, cytoplasts of interphase zygotes, after the removal of pronuclei, have been incapable to support such development (10). When the pronuclear membrane and chromatin of interphase zygotes are selectively removed but other pronuclear components, including the nucleolus, are left in the cytoplasm, these cytoplasts support development of SCNT embryos (11). These results suggested that an undefined activity or material(s) in pronuclei of zygotes facilitates the reprogramming of transferred somatic cell nuclei (9, 11). To determine whether the maternal nucleolus alone facilitates the somatic cell reprogramming in zygotic cytoplasts, we cotransferred the cumulus cell nuclei and the isolated oocyte nucleoli into cytoplasts from interphase enucleated mouse zygotes (Zygote-SCNT). When the somatic cell nuclei were transferred into the cytoplast of M-phase enucleated zygotes, the resulting embryos assembled visible nucleoli (fig. S9, M-phase Zygote-SCNT) and developed to the blastocyst stage (table S4). In contrast, the embryos from Zygote-SCNT cleaved at 36 hours after human chorionic gonadotropin injection, but they arrested at the two-cell stage or fragmented, and they never developed to blastocysts. These two-cell embryos derived from zygote-SCNT never assembled nucleoli in their blastomere nuclei (fig. S9, Zygote-SCNT). Coinjection of oocyte nucleoli along with somatic cell nuclei resulted in two-cell–stage embryos with prominent nucleoli in their nuclei (fig. S9, Zygote-SCNT+Nucleolus). This nucleolar coinjection, however, did not rescue embryonic development, and embryos were typically arrested between the two- and four-cell stages (table S4). This indicates that, in addition to the nucleolar material, some other
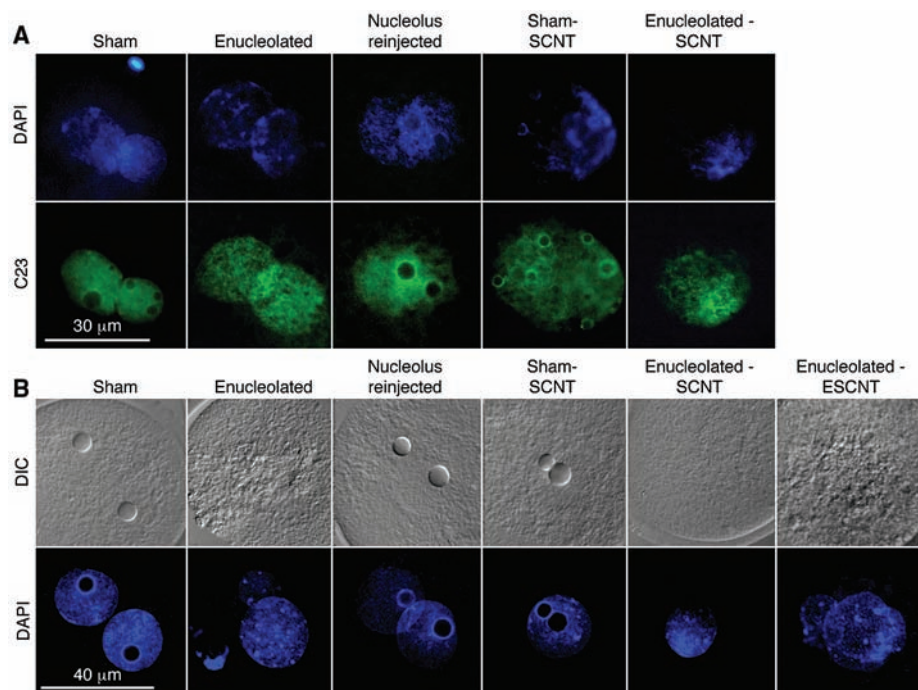
Fig. 2. The presence of nucleoli in pronuclei of activated enucleolated oocytes after reinjection of isolated oocyte nucleoli or after SCNT. No nucleolus was formed in the embryos originating from enucleolated (Enucleolated) oocytes, from somatic cell (cumulus cell) nucleus–transferred enucleolated (Enucleolated-SCNT) oocytes, and from ES cell nucleus–transferred enucleolated (Enucleolated-ESCNT) oocytes. Nucleoli in pronuclei were detected by immunofluorescence labeling against C23 (green) in the pig (**A**) and by DIC microscopy in the mouse (**B**). DAPI staining marks chromatin in blue.

Table 1. In vitro development of pig parthenotes and SCNT embryos derived from enucleolated oocytes. The values in parentheses represent the percentages of oocytes cleaved at 48 hours and forming blastocysts at 6 days after activation.

	Type of pig oocyte				
	Sham	Enucleolated	Nucleolus reinjected	Sham-SCNT	Enucleolated-SCNT
Total number of oocytes examined	148	228	194	156	200
Total number and % of oocytes cleaved	128 (86)	168 (74)	174 (90)	120 (77)	175 (88)
Total number and % of oocytes forming blastocysts	92 (62)	0 (0)	72 (37)	48 (31)	0 (0)
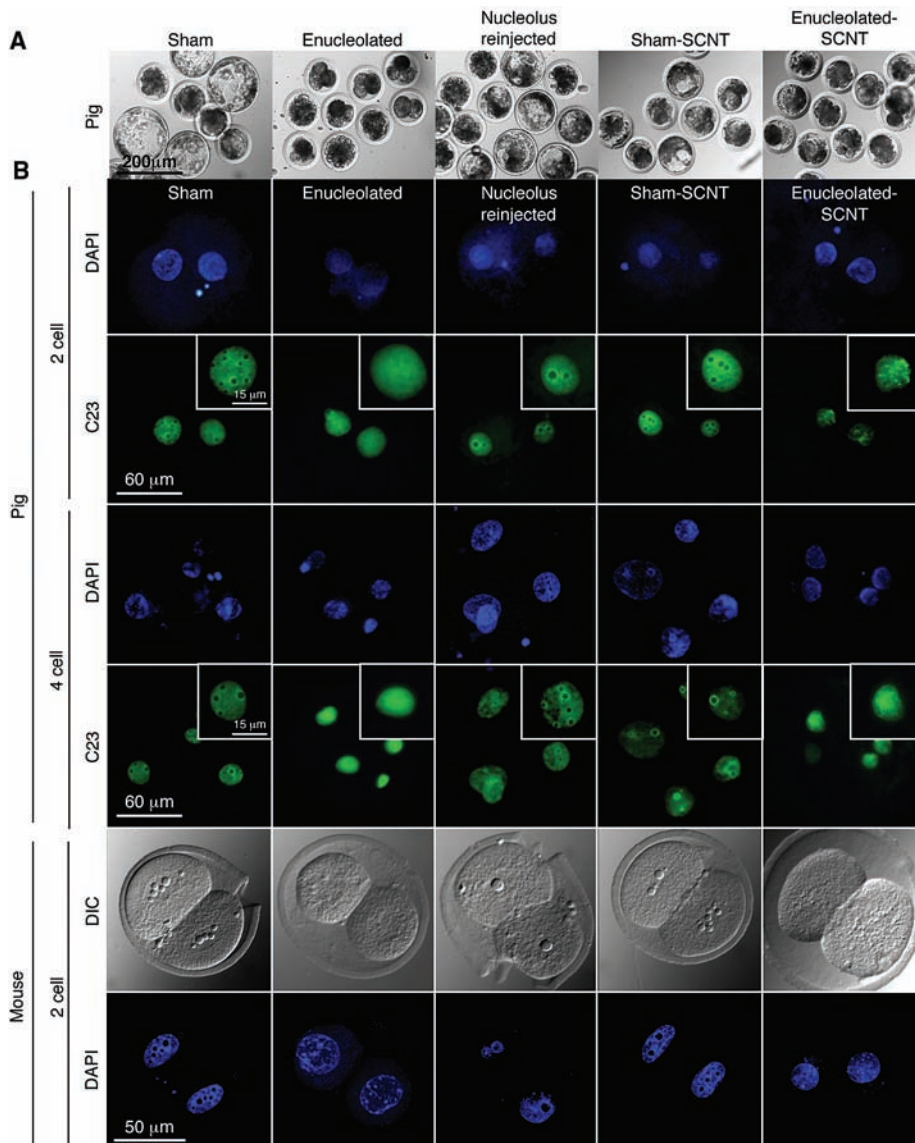
Fig. 3. Requirement of the oocyte nucleolus for early embryonic development. (**A**) Embryos originating from enucleolated and from enucleolated-SCNT oocytes, which were lacking the oocyte nucleolar material, were fragmented or arrested at the four- to eight-cell stage 6 days after electroactivation, whereas the embryos with nucleoli developed to blastocysts. (**B**) The absence of nucleoli in the embryos originating from enucleolated and enucleolated-SCNT oocytes. Nucleoli were visualized by immunofluorescence labeling against C23 (green) in the pig and by DIC microscopy in the mouse. Insets show nuclei from pig embryos at higher magnification. DAPI staining marks chromatin in blue.

nucleoplasmic components are essential for the successful development of zygote-SCNT embryos.

Our experiments demonstrate that nucleoli in mammalian zygotes are exclusively maternally inherited via the oocyte. We have also showed the necessity of the entire nucleolar complement for successful embryonic development. Because the nucleolus in the spermatozoon is eliminated during spermiogenesis (12–14), the oocyte nucleolar material is essential for the reassembly of newly formed nucleoli in both female and male pronuclei. We have found that after SCNT or ES cell

nuclear transfer, the nucleolus apparently originates from the maternal source (oocyte), but we cannot exclude a possible minor contribution from the somatic and ES cells. This may be one of the reasons why cytoplasts of oocytes after germinal vesicle breakdown, as well as cytoplasts of mitotic zygotes, support development after nuclear transfer, whereas cytoplasts of oocytes before germinal vesicle breakdown, as well as cytoplasts of interphase zygotes, do not (9, 15). The substances that are stored either in germinal vesicles or in pronuclei and that are essential for the reprogramming of somatic cell nuclei are not

yet defined, but our results document that one of them is the nucleolus.

After fertilization, embryos drive their development for one or several cell cycles using the stockpile of molecules originating from oocytes. Therefore, it is speculated that early protein synthesis uses maternally stored ribosomes and RNAs. Functional nucleoli are assembled later on, when the embryonic genome becomes activated (5, 6, 16, 17). The nucleolar materials originating from oocytes may represent the precursor molecules required for the assembly of fully functional nucleoli at a later stage, as the embryo develops (16). Our results suggest that the oocyte nucleolus provides the materials to build nucleoli not only in pronuclei of zygotes but also in pseudo-pronuclei formed after SCNT. Thus, the oocyte nucleolar material is essential for successful early embryonic development in mammals.

References and Notes
1. C. A. Hutchison III, J. E. Newbold, S. S. Potter, M. H. Edgell, *Nature* **251**, 536 (1974).
2. G. Schatten, *Dev. Biol.* **165**, 299 (1994).
3. P. Sutovsky, G. Schatten, *Reprod. Nutr. Dev.* **38**, 629 (1998).
4. N. Crozet, J. Motlik, D. Szollosi, *Biol. Cell* **41**, 35 (1981).
5. V. Kopecny, *Reprod. Nutr. Dev.* **29**, 589 (1989).
6. P. Hyttel *et al.*, *Reproduction* **122**, 21 (2001).
7. B. Bjerregaard, P. Maddox-Hyttel, *Anim. Reprod. Sci.* **82-83**, 605 (2004).
8. J. Fulka Jr., R. M. Moor, P. Loi, J. Fulka, *Theriogenology* **59**, 1879 (2003).
9. D. Egli, J. Rosains, G. Birkhoff, K. Eggan, *Nature* **447**, 679 (2007).
10. J. McGrath, D. Solter, *Science* **226**, 1317 (1984).
11. P. Greda, J. Karasiewicz, J. A. Modlinski, *Reproduction* **132**, 741 (2006).
12. V. Monesi, *Exp. Cell Res.* **39**, 197 (1965).
13. M. C. Schultz, C. P. Leblond, *Am. J. Anat.* **189**, 1 (1990).
14. I. K. Takeuchi, Y. K. Takeuchi, *J. Struct. Biol.* **103**, 104 (1990).
15. S. Gao *et al.*, *Biol. Reprod.* **67**, 928 (2002).
16. J. Tesarik, V. Kopecny, M. Plachot, J. Mandelbaum, *Development* **101**, 777 (1987).
17. D. Viuff *et al.*, *Biol. Reprod.* **66**, 629 (2002).
18. We thank M. M. Matzuk (at Baylor College of Medicine) for his critical discussion and for providing the NPM2 antibody; U. Kikkawa (at Kobe University) and K. Yamanaka, S. Wakayama, and T. Wakayama (at Center for Developmental Biology RIKEN) for their generous support; and the staff of the Kobe Meat Inspection Office for supplying the pig ovaries. This work was supported by Grants-in-Aid from the Ministry of Education, Culture, Sports, Science, and Technology of Japan (to S.O., M.S., and T.M.) and by grants MZe 0002701401 and GACR STE 05/E004 (to J.F. Jr. and H.F.).

Profiling Essential Genes in Human Mammary Cells by Multiplex RNAi Screening

Jose M. Silva,[1] Krista Marran,[1] Joel S. Parker,[3] Javier Silva,[1] Michael Golding,[1] Michael R. Schlabach,[2] Stephen J. Elledge,[2] Gregory J. Hannon,[1]* Kenneth Chang[1]

By virtue of their accumulated genetic alterations, tumor cells may acquire vulnerabilities that create opportunities for therapeutic intervention. We have devised a massively parallel strategy for screening short hairpin RNA (shRNA) collections for stable loss-of-function phenotypes. We assayed from 6000 to 20,000 shRNAs simultaneously to identify genes important for the proliferation and survival of five cell lines derived from human mammary tissue. Lethal shRNAs common to these cell lines targeted many known cell-cycle regulatory networks. Cell line–specific sensitivities to suppression of protein complexes and biological pathways also emerged, and these could be validated by RNA interference (RNAi) and pharmacologically. These studies establish a practical platform for genome-scale screening of complex phenotypes in mammalian cells and demonstrate that RNAi can be used to expose genotype-specific sensitivities.

The observation of genetic interactions is key to the definition of cellular networks. RNAi has enabled genetic approaches in both cultured mammalian cells (1–5) and intact animals (6–9). Large-scale screens of small interfering RNA (siRNA) (10–12) and shRNA collections (5, 13–16) have generally adopted a one-by-one approach, interrogating phenotypes in a well-based format. This requires both considerable infrastructure and a substantial investment for each cell line to be screened. Alternatively, shRNA collections can be screened by assaying enrichment from pools, but this limits the range of phenotypes that can be addressed. Our focus was identifying essential genes or synthetically lethal genetic interactions through shRNAs that were selectively depleted from populations. This type of screen holds promise for the discovery of novel targets for cancer therapy and genetically validated combination therapies. Previously, one such screen was reported; however, this tested only ~500 shRNAs in a single pool (17). We therefore sought methods that allow multiplex analysis of phenotypic outputs on a genomic scale.

Pooled libraries drew from our previous collections wherein shRNAs are carried in a backbone derived from miR-30 (18). Combining RNA polymerase II promoters with miR-30–based shRNAs permits efficient suppression even with a single-copy integrant (19, 20). Therefore, pooled shRNAs were transferred from pSM2 (18) to pLMP (19), wherein shRNA expression is driven from the murine stem cell virus long-terminal repeat promoter. Three different pools, containing ~6000, ~10,000, and ~20,000 shRNAs, were constructed to test screening at varying scales and levels of population complexity. Target cell populations were infected such that each cell contained, on average, a single integrated virus, and each individual shRNA occupied ~1000 cells. Three parallel infections generated biological replicate samples. Because our goal was to identify essential genes, genomic DNA was prepared from each replicate at three time points during a simple outgrowth assay (Fig. 1A).

Each shRNA cassette contains two unique identifiers: the shRNA itself and a random 60-nucleotide barcode. Barcode sequences were determined for the human shRNA library, and custom, multiplex format microarrays were prepared that contained both barcode and half-hairpin (HH) probes (21) (Fig. 1B). Proviral DNA fragments encompassing both shRNAs and barcodes were amplified from genomic DNA pools and hybridized to arrays in competition with a common reference.

We established a rigorous data analysis pipeline (22) for analyzing pooled shRNA screens. Correlations between biological replicates were high but diminished at later time points, whereas correlations between the reference channels remained unchanged (table S1). Overall, a gene was scored as a candidate if either its barcode or shRNA probe showed greater than 2-fold change with a false discovery rate (FDR) <10%.

We began with a pooled analysis of 6000 (6K) shRNAs in MCF-10A and MDA-MB-435. Although enriched gene sets varied considerably, similar numbers and largely overlapping gene sets showed depletion in both cell lines (tables S2 and S3). Among negatively selected shRNAs were many targeting regulators of the cell division cycle (23, 24) (table S3). These included cyclins, cell division cycle (CDC) proteins, E2F family members, minichromosome maintenance deficient genes, proliferating cell nuclear antigen, and RNA polymerase II–associated genes. Additionally, the proteasome (15 of 25 subunits; $P = 5.61 \times 10^{-5}$) and anaphase-promoting complex/cyclosome (APC/C; 6 of 11 subunits; $P = 0.0139$) scored as being essential in both cell lines (table S3).

To validate candidates, we constructed a regulated shRNA vector, which linked shRNA and green fluorescent protein expression (fig. S1A). Inducible shRNAs against two APC/C subunits,

[1]Watson School of Biological Sciences, Howard Hughes Medical Institute, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY 11724, USA. [2]Department of Genetics, Center for Genetics and Genomics, Brigham and Women's Hospital, Howard Hughes Medical Institute, Harvard Medical School, Boston, MA 02115, USA. [3]Expression Analysis, Inc., 4324 South Alston Avenue, Durham, NC 27713, USA.

*To whom correspondence should be addressed. E-mail: hannon@cshl.edu

Fig. 1. Experimental approach. (**A**) shRNA plasmids were packaged into retroviruses in triplicate and introduced into replicate target cell populations at a multiplicity of ~0.3 to achieve ~1 integrant per cell. Over a 2-week culture period, time points were collected on day 2 or day 4 after infection and then once each week for 2 weeks. (**B**) The shRNA guide strand and the barcode region were amplified from genomic DNA from screening pools. Polymerase chain reaction (PCR) products were gel-purified, labeled, and hybridized to multiplex arrays in competition with a common reference.
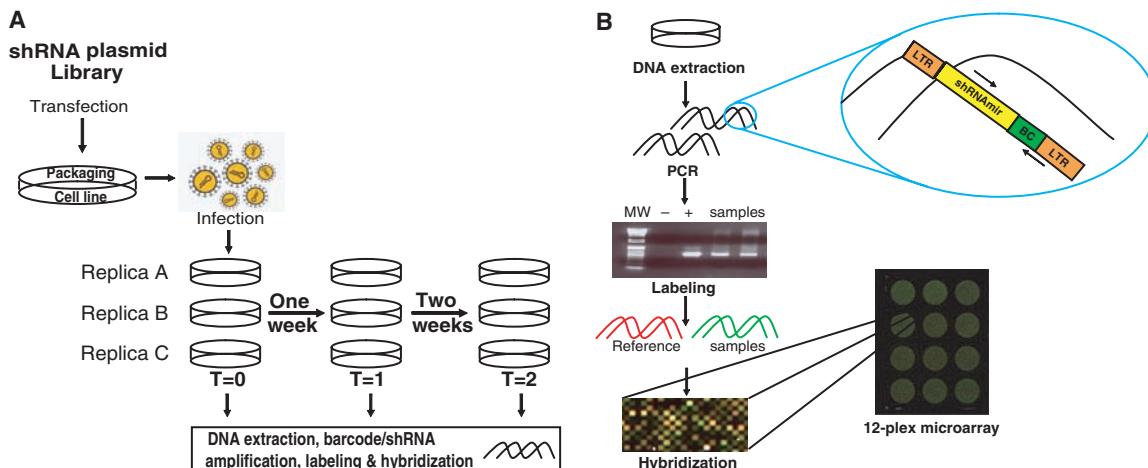
ANAPC2 and 4, inhibited the growth of MCF-10A in a manner that correlated with mRNA knockdown (fig. S1A and Fig. 2A). Similarly, MDA-MB-435 was sensitive to ANAPC2 depletion (Fig. 2C). Nineteen additional MCF-10A lines were constructed with inducible shRNAs targeting 11 different candidates (Fig. 2B). Of these, 16 lines exhibited shRNA-dependent growth inhibition (30% to 95%), which correlated with mRNA knockdown in 14 cases. The exceptions were CDC-5L and DKC-1, where growth suppression could be due to off-target effects (Fig. 2B).

Among additional candidates were MAD2 and BUBR1, mitotic checkpoint proteins required for regulation of sister chromatid separation (24–26), and Kinesin-7/CENP-E, a component of the kinetochore (27). MAD2/MAD2L1 and Kinesin-7/CENP-E were validated as being essential in MCF-10A (table S3 and Fig. 2B). CENP-E depletion also inhibited growth in MDA-MB-435 (table S3 and Fig. 2C). Considered together, these studies showed that multiplex RNAi screens successfully identified essential components of cell growth and survival networks.

We also screened higher complexity populations containing 10,000 (10K) or 20,000 (20K) shRNAs. The 10K pool was introduced into MDA-MB-231, T-47D and ZR-75-1. The most complex pool (20K) was introduced into MCF-10A to allow direct comparison with the 6K screen. In all cases, cell numbers were scaled to maintain a representation of 1000 cells per shRNA. The quality of each screen was similar, with high correlations between biological replicates (table S4). We assessed the consistency of the MCF-10A screens by comparing depleted gene sets for the 6K and 20K pools. FDR thresholds were the same for both data sets ($q <$ 0.1), but the fold-change criterion was relaxed from 2-fold to 1.5-fold for the 20K screen so that similar numbers of candidates were compared. A set of 172 genes ($P = 1.123 \times 10^{-9}$) overlapped in both data sets, despite some differences in the protocols

Fig. 2. Validation of genes essential to multiple cell lines. Cell viability assays (bars) were performed on cell lines (MCF-10A or MDA-MB-435) expressing individual candidate shRNAs. Tables below the graphs show the level of target suppression, determined by quantitative real-time fluorescence PCR, with or without shRNA induction (indicated). An shRNA targeting luciferase (FF) and no shRNA serve as negative controls. (**A**) APC subunits ANAPC2 and ANAPC4 were suppressed by multiple hairpins in MCF-10A (1 to 5 for ANAPC4 and 1 to 4 for ANAPC2). Cell viability assays were carried out for 7 days after shRNA induction. (**B**) Nineteen additional inducible MCF-10A cell lines were generated to validate shRNAs that were depleted in the screen. Viability assays were carried out for 11 days. (**C**) Validation of shRNAs that were depleted (left of the yellow line) or not depleted (right of the yellow line) in MDA-MB-435 cells. Viability assays were carried out for 11 days. (**D**) The Venn diagram illustrates the 166 depleted genes that were common to both screens and the 35 and 3 genes specifically depleted from MCF-10A and MDA-MB-435, respectively. shRNAs targeting P-TEFb components CDK9 and cyclin T2 were both depleted specifically in MCF-10A. The graph shows a dose-response curve for growth inhibition of MCF-10A and MDA-MB-435 cells by DRB over a 72-hour period.
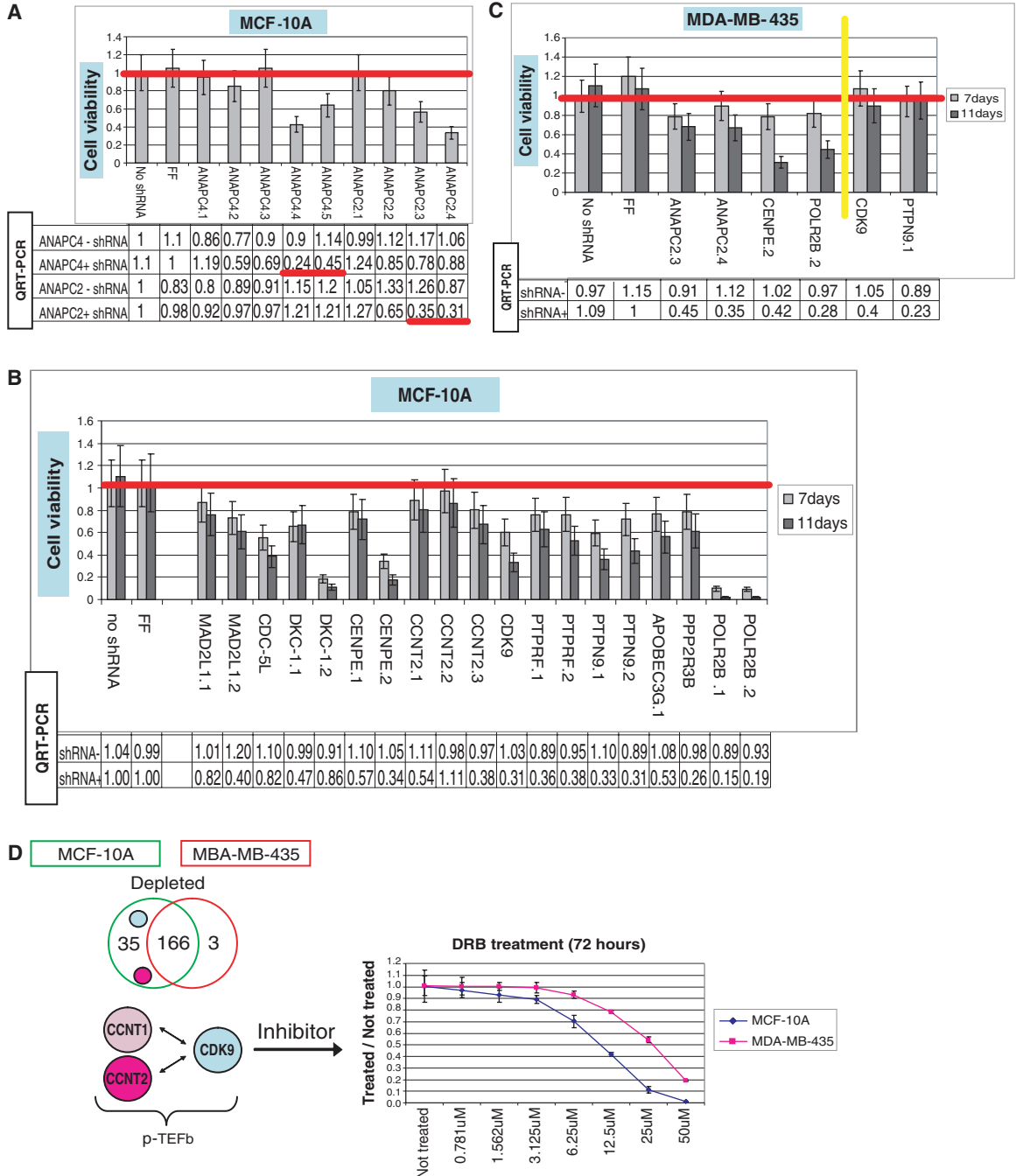
used to carry out each screen, and most of the validated targets from the 6K screen were found in the overlapping list of essential genes (tables S5 and S6). This suggests that a pool of ~20K shRNAs can be effectively screened.

We next sought to uncover cell line–specific genetic sensitivities that might reflect differences in the genetic constitutions of MCF-10A, MDA-MB-231, MDA-MB-435, and ZR-75.1. Initial comparisons focused on the 6K screens done with MCF-10A and MDA-MB-435. Filtering for shRNAs that had a low FDR ($q < 0.1$) and at least 2-fold depletion in MCF-10A but no more than 1.2-fold depletion in MDA-MB-435 yielded 35 genes (table S7). This compares to 166 genes that were important for growth in both cell lines and 3 genes that were differentially required in MDA-MB-435 (Fig. 2D and table S7). Among the candidates required

in MCF-10A were two components of P-TEFb, CDK9 and cyclin T2 (28). We verified this differential sensitivity using both conditional shRNA expression (Fig. 2, B and C) and pharmacological inhibition (Fig. 2D). CDK9 is a DRB-sensitive kinase (28). Although DRB may also target other proteins, MCF-10A showed greater sensitivity to its effects than MDA-MB-435 cells (Fig. 2D).

We repeated the 6K screens of MCF-10A and MDA-MB-435 cells on the same array platform as the 10K screens of MDA-MB-231, T-47D, and ZR-75.1 cells and integrated the results (table S8). This was possible because more than 90% of the 6K shRNA set was contained within the 10K pool. Clustering of the resulting sensitivities (i.e., by fold-change, considering only shRNAs with $q < 0.1$) yielded a dendrogram wherein the more normal MCF-10A segregated from the other, more trans-

formed lines (Fig. 3A). MDA-MB-435 also segregated, perhaps reflecting the observation that it is more related, by expression profiling, to melanoma than breast epithelia. Finally, the remaining lines separated into a group containing T-47D and ZR-75.1, both luminal tumor cell lines, and MDA-MB-231, a basal tumor cell line (29) (Fig. 3A).

Viewing this portrait of shRNA sensitivity in more detail revealed a number of pathways and complexes that were differentially required in MCF-10A. These included epidermal growth factor receptor (EGFR), an effect that could be reproduced pharmacologically using the EGFR inhibitor Tarceva (30) (Fig. 3B). DNA methyltransferases also scored either above or close to the threshold (table S8 and Fig. 3B). In accord with these results, MCF-10A cells showed a more than 50-fold greater sensitivity to 5-aza-deoxycytidine, a methyltrans-
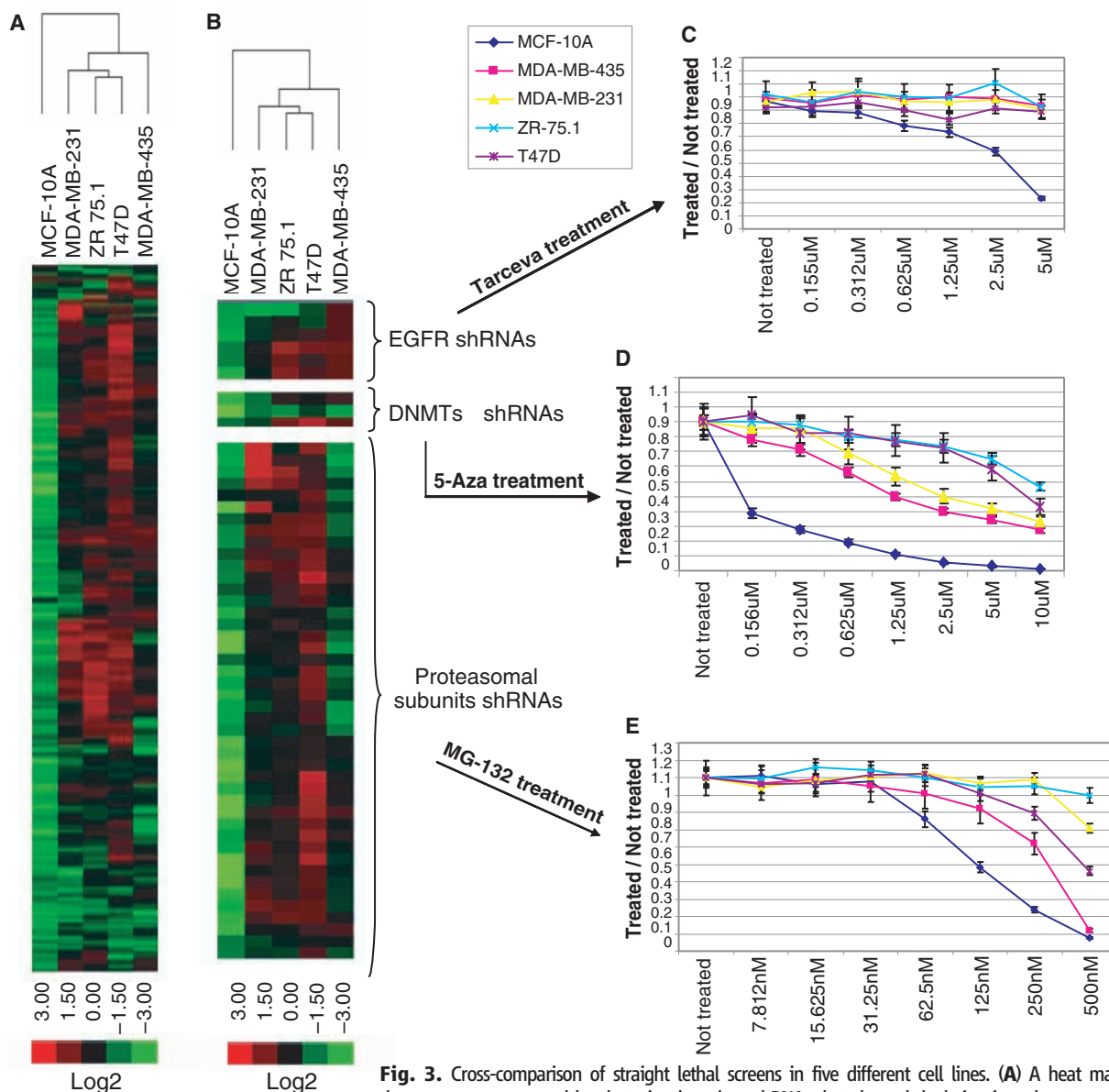
Fig. 3. Cross-comparison of straight lethal screens in five different cell lines. (**A**) A heat map and dendrogram were generated by clustering based on shRNAs that showed depletion in at least one cell line. (**B**) shRNAs targeting selected complexes or pathways were chosen from (**A**) to highlight responses to EGFR, methyltransferase, and proteasome lesions. Pharmacological inhibition of these pathways by treatment with Tarceva (**C**), 5-aza-deoxycytidine (**D**), or MG-132 (**E**) was used to validate the predictions of the shRNA screen.

ferase suicide substrate (*31*), than the other cell lines. As a final example, numerous proteasome subunits were preferentially depleted from MCF-10A (table S8 and Fig. 3B). These cells showed the greatest sensitivity to a proteasome inhibitor, MG-132 (*32*). Interestingly, MDA-MB-435 showed an intermediate level of sensitivity to the drug, and this was reflected precisely in their intermediate level of depletion of proteasomal shRNAs during the screen (table S8 and Fig. 3B).

We have validated a highly scalable approach for screening shRNA libraries. Although we used a phenotypic filter reflecting growth and survival, virtually any characteristic that allows separation of phenotypically distinct cells can be applied. We also validated the ability of functional shRNA screening to separate cell lines based on their genetic vulnerabilities in a manner that reflects their already defined characteristics (e.g., immortal versus tumor, basal versus luminal). Although one could attribute selective dependency to culture conditions in some cases, the overwhelming concordance of the shRNAs that affect proliferation and survival across these lines, many of which are cultured identically, strongly argues against this being a pervasive explanation. In all, this approach enables genome-wide screens for tumor-specific vulnerabilities to be carried out on large numbers of tumor lines. Moreover, it permits rational searches for lesions that synergize with existing therapeutics to produce a path toward genetically informed combination therapies.

References and Notes

1. P. J. Paddison, A. A. Caudy, G. J. Hannon, *Proc. Natl. Acad. Sci. U.S.A.* **99**, 1443 (2002).
2. S. M. Elbashir *et al.*, *Nature* **411**, 494 (2001).
3. L. Scherer, J. J. Rossi, *Curr. Pharm. Biotechnol.* **5**, 355 (2004).
4. A. Friedman, N. Perrimon, *Curr. Opin. Genet. Dev.* **14**, 470 (2004).
5. R. Bernards, T. R. Brummelkamp, R. L. Beijersbergen, *Nat. Methods* **3**, 701 (2006).
6. M. T. Hemann *et al.*, *Nat. Genet.* **33**, 396 (2003).
7. O. Snove Jr., J. J. Rossi, *Nat. Methods* **3**, 689 (2006).
8. M. H. Farah, *Curr. Drug Deliv.* **4**, 161 (2007).
9. A. de Fougerolles, H. P. Vornlocher, J. Maraganore, J. Lieberman, *Nat. Rev. Drug Discov.* **6**, 443 (2007).
10. P. Aza-Blanc *et al.*, *Mol. Cell* **12**, 627 (2003).
11. A. W. Whitehurst *et al.*, *Nature* **446**, 815 (2007).
12. F. Natt, *Curr. Opin. Mol. Ther.* **9**, 242 (2007).
13. P. J. Paddison *et al.*, *Nature* **428**, 427 (2004).
14. T. F. Westbrook *et al.*, *Cell* **121**, 837 (2005).
15. D. E. Root *et al.*, *Nat. Methods* **3**, 715 (2006).
16. C. Gazin *et al.*, *Nature* **449**, 1073 (2007).
17. V. N. Ngo *et al.*, *Nature* **441**, 106 (2006).
18. J. M. Silva *et al.*, *Nat. Genet.* **37**, 1281 (2005).
19. R. A. Dickins *et al.*, *Nat. Genet.* **37**, 1289 (2005).
20. F. Stegmeier, G. Hu, R. J. Rickles, G. J. Hannon, S. J. Elledge, *Proc. Natl. Acad. Sci. U.S.A.* **102**, 13212 (2005).
21. M. R. Schlabach *et al.*, *Science* **319**, 620 (2008).
22. Materials and methods are available as supporting material on *Science* Online.
23. K. I. Nakayama, K. Nakayama, *Nat. Rev. Cancer* **6**, 369 (2006).
24. J. M. Peters, *Nat. Rev. Mol. Cell Biol.* **7**, 644 (2006).
25. G. Fang, *Mol. Biol. Cell* **13**, 755 (2002).
26. G. J. Kops, D. R. Foltz, D. W. Cleveland, *Proc. Natl. Acad. Sci. U.S.A.* **101**, 8699 (2004).
27. T. J. Yen *et al.*, *Nature* **359**, 536 (1992).
28. J. Peng, Y. Zhu, J. T. Milton, D. H. Price, *Genes Dev.* **12**, 755 (1998).
29. R. M. Neve *et al.*, *Cancer Cell* **10**, 515 (2006).
30. J. D. Moyer *et al.*, *Cancer Res.* **57**, 4838 (1997).
31. L. Jackson-Grusby *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **94**, 4681 (1997).
32. V. J. Palombella *et al.*, *Cell* **78**, 773 (1994).
33. We thank members of the Hannon and Lowe laboratories for helpful discussion and T. Moore of Open Biosystems for help and support. This work was supported by grants from NIH and the Department of Defense to G.J.H. and J.M.S. and by a kind gift from Kathryn W. Davis. G.J.H. and S.J.E. have a paid consulting relationship with Open Biosystems.

Cancer Proliferation Gene Discovery Through Functional Genomics

Michael R. Schlabach,[1]* Ji Luo,[1]* Nicole L. Solimini,[1]* Guang Hu,[1]* Qikai Xu,[1] Mamie Z. Li,[1] Zhenming Zhao,[1] Agata Smogorzewska,[1,2] Mathew E. Sowa,[3] Xiaolu L. Ang,[3] Thomas F. Westbrook,[1] Anthony C. Liang,[1] Kenneth Chang,[4] Jennifer A. Hackett,[1] J. Wade Harper,[3] Gregory J. Hannon,[4] Stephen J. Elledge[1]†

Retroviral short hairpin RNA (shRNA)–mediated genetic screens in mammalian cells are powerful tools for discovering loss-of-function phenotypes. We describe a highly parallel multiplex methodology for screening large pools of shRNAs using half-hairpin barcodes for microarray deconvolution. We carried out dropout screens for shRNAs that affect cell proliferation and viability in cancer cells and normal cells. We identified many shRNAs to be antiproliferative that target core cellular processes, such as the cell cycle and protein translation, in all cells examined. Moreover, we identified genes that are selectively required for proliferation and survival in different cell lines. Our platform enables rapid and cost-effective genome-wide screens to identify cancer proliferation and survival genes for target discovery. Such efforts are complementary to the Cancer Genome Atlas and provide an alternative functional view of cancer cells.

We have recently generated barcoded, microRNA-based shRNA libraries targeting the entire human genome that can be expressed efficiently from retroviral or lentiviral vectors in a variety of cell types for stable gene knockdown (*1*, *2*). Furthermore, we have also developed a method of screening complex pools of shRNAs using barcodes coupled with microarray deconvolution to take advantage of the highly parallel format, low cost, and flexibility in assay design of this approach (*2*, *3*). Although barcodes are not essential for enrichment screens (positive selection) (*3–5*), they are critical for dropout screens (negative selection) such as those designed to identify cell-lethal or drug-sensitive shRNAs (*6*). Hairpins that are depleted over time can be identified through the competitive hybridization of barcodes derived from the shRNA population before and after selection to a microarray (Fig. 1A).

We previously described the use of 60-nucleotide barcodes for pool deconvolution (*2*, *3*). To provide an alternative to these barcodes that enables a more rapid construction and screening of shRNA libraries, we have developed a methodology called half-hairpin (HH) barcoding for deconvoluting pooled shRNAs (*7*). We took advantage of the large 19-nucleotide hairpin loop of our mir30-based platform and designed a polymerase chain reaction (PCR) strategy that amplifies only the 3′ half of the shRNA stem (Fig. 1B). As compared with full-hairpin sequences for microarray hybridization (*8*, *9*), HH barcodes entirely eliminate probe

self-annealing during microarray hybridization (Fig. 1C and fig. S1, A and B), providing the critical dynamic range necessary for pool-based dropout screens. HH barcode signals are highly reproducible in replicate PCRs ($R = 0.973$, fig. S1A), highly specific (0.5% cross-reaction) (fig. S1C), and display a reasonable, although slightly compressed, dynamic range in mixing experiments with varied subpool inputs that are quantified by microarray hybridization (fig. S1, D and E). Taken together, these results indicate that HH barcodes are alternatives to the 60-nucleotide barcodes originally designed into our library.

Our central goal is to develop the means to rapidly perform dropout screens to systematically identify genes required for cancer cell proliferation and survival that could represent new drug targets. We used our screening platform to interrogate human DLD-1 and HCT116 colon cancer cells, human HCC1954 breast cancer cells, and normal human mammary epithelial cells (HMECs). We compared colon and breast cancer cells—two types of cancers with distinct origins—to maximize our ability to identify common and cancer-specific growth regulatory pathways. Recent large-scale efforts have

[1]Howard Hughes Medical Institute and Department of Genetics, Center for Genetics and Genomics, Brigham and Women's Hospital, Harvard Medical School, Boston, MA 02115, USA. [2]Department of Pathology, Massachusetts General Hospital (MGH), Boston, MA 02114, USA. [3]Department of Pathology, Harvard Medical School, Boston, MA 02115, USA. [4]Watson School of Biological Sciences, Howard Hughes Medical Institute, Cold Spring Harbor Laboratory, 1 Bungtown Road, Cold Spring Harbor, NY 11724, USA.

*These authors contributed equally to this work.
†To whom correspondence should be addressed. E-mail: selledge@genetics.med.harvard.edu

identified a distinct spectrum of mutations in these two cancer types (*10*, *11*). Also, the comparison between cancer cells and normal cells should reveal potential growth and survival adaptations specific to cancer cells. We constructed a highly complex pool of 8203 distinct shRNAs targeting 2924 genes consisting of annotated kinase, phosphatase, ubiquitination pathway, and cancer-related genes (table S1). We chose these genes because they are central regulators of signaling pathways that should provide a rich source of phenotypic perturbation. These

shRNAs were placed in a murine stem cell virus (MSCV)–derived retroviral vector (*12*), MSCV-PM, that functions efficiently at single copy.

We screened each cell line in independent triplicates (*7*). Cells were infected with an average representation of 1000 independent integrations per shRNA and with a multiplicity of infection of 1 to 2. Initial reference samples were collected 48 to 72 hours after infection. The remaining cells were puromycin-selected, propagated for several weeks, and collected again as the end samples. HH bar-

codes were PCR-recovered from genomic DNA, labeled with Cy5 and Cy3 dyes, respectively, and hybridized to a HH barcode microarray (Fig. 1A). The Cy3/Cy5 signal ratio of each probe reports the change in relative abundance of a particular shRNA between the beginning and the end of the experiment. Correlations between initial samples across the triplicates and between the initial and end samples within each replica were high, indicating that the triplicates were highly reproducible and representation was well maintained throughout the experiment.
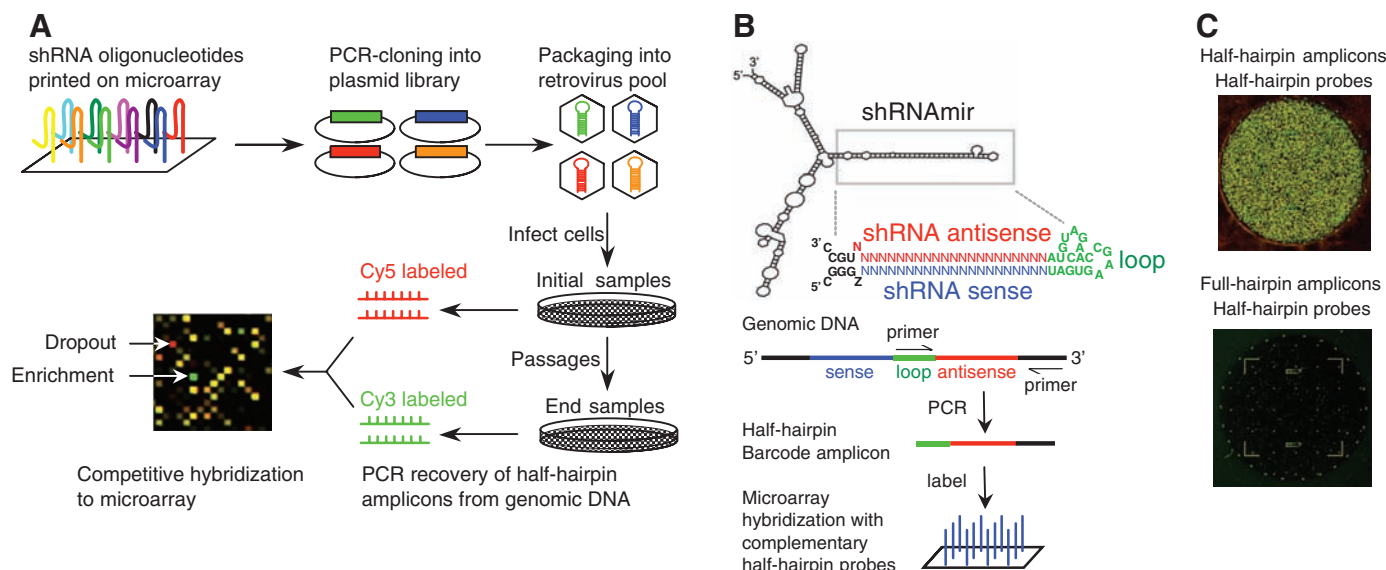
Fig. 1. Overview of the pool-based dropout screen with barcode microarrays. (**A**) Schematic of library construction and screening protocol. (**B**) Schematic of the HH barcode hybridization. (**C**) Comparison between HH amplicons (top) and full-hairpin PCR amplicons (bottom) on an HH probe microarray.
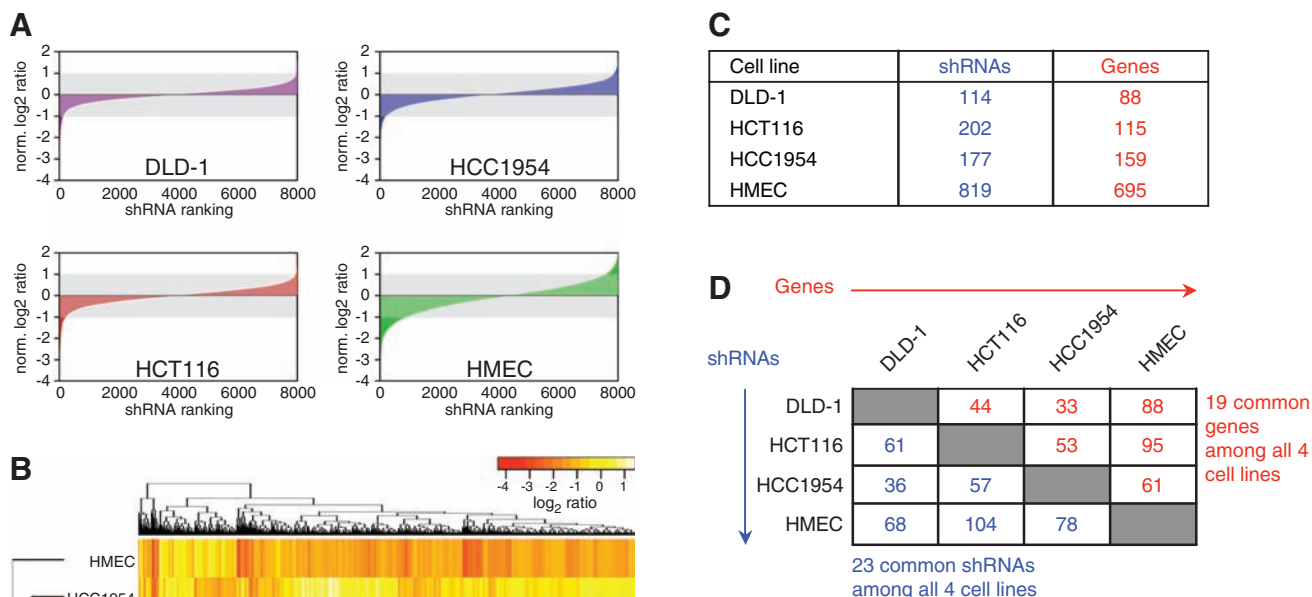
Fig. 2. Pool-based dropout screen for genes required for cancer cell viability. (**A**) Overview of shRNA pool behavior in the screen. For each cell line, shRNAs were ranked on the basis of their mean normalized $\log_2$ Cy3/Cy5 ratios. The shaded rectangle indicates the $\log_2$ ratio range within which an shRNA's abundance was considered unchanged. (**B**) Clustering of the four cell lines with the antiproliferative shRNAs identified in the screen. The color scale represents mean normalized $\log_2$ Cy3/Cy5 ratios of the probes. (**C**) Antiproliferative shRNAs and genes that scored in the screen for each cell line are shown. (**D**) Summary of the common shRNAs (blue) and genes (red) identified in the screen. Overlapping antiproliferative shRNAs/genes between pairwise combinations of cell lines are displayed (DLD-1 and HMEC have more overlapping genes than shRNAs because, in some cases, different sets of shRNAs targeting the same gene scored in each line).

REPORTS

To identify shRNAs that consistently changed in abundance in each cell line, we analyzed data sets using a custom statistical package based on the Linear Models for Microarray data (Limma) method (13) for two-color cDNA microarray analysis (7). Whereas most shRNAs showed little changes in their abundance over time (log$_2$ ratio between –1 and 1), a small fraction of shRNAs showed depletion (Fig. 2A). Based on their shRNA dropout signatures, unsupervised hierarchical clustering segregated the three cancer cell lines from the normal HMECs, likely reflecting fundamental differences between cancer cells and normal cells (Fig. 2B). Furthermore, the two colon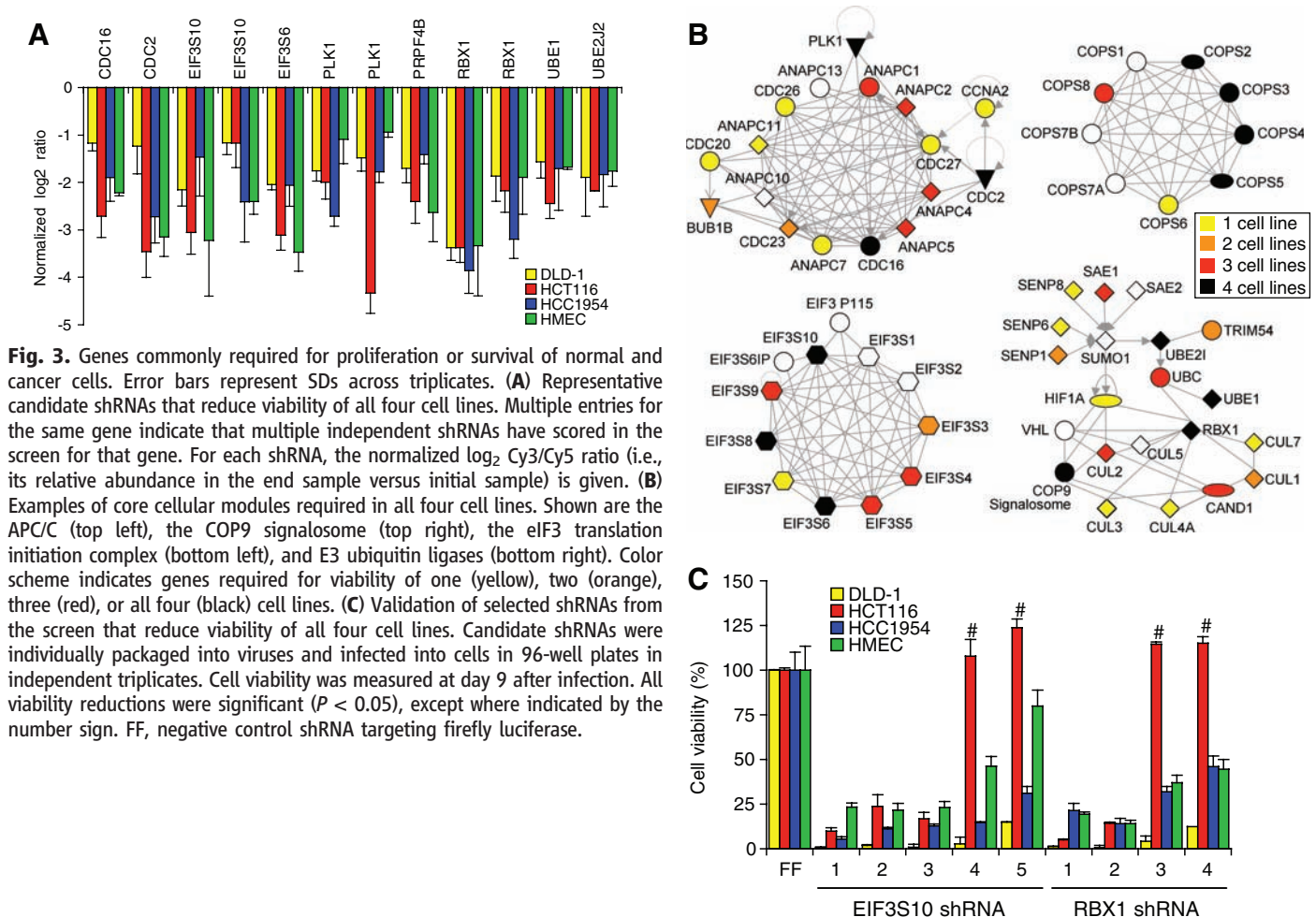 cancer cell lines were more similar to each other than to the breast cancer line, reflecting the differences in their tissues of origin and paths to tumorigenesis. Overall, we found 114 shRNAs (1.4%) representing 88 genes (3.0%) in DLD-1 cells, 202 shRNAs (2.5%) representing 115 genes (3.9%) in HCT116 cells, 177 shRNAs (2.2%) representing 159 genes (5.4%) in HCC1954 cells, and 819 shRNAs (10.0%) representing 695 genes (23.8%) in HMECs showed statistically significant depletion (Fig. 2C and tables S3 to S6). The lists of antiproliferative shRNAs show significant overlap ($P < 1 \times 10^{-40}$), with 23 shRNAs and 19 genes scoring in all four lines (Fig. 2D). As expected, our screen recovered components of core cellular modules essential for all

cell lines (Fig. 3, A and B). For example, shRNAs against multiple subunits of the anaphase promoting complex/cyclosome (APC/C) (DLD-1, $P = 9.65 \times 10^{-5}$; HCT116, $P = 2.99 \times 10^{-9}$; HCC1954, $P = 1.41 \times 10^{-5}$; HMEC, $P = 5.80 \times 10^{-6}$), the COP9 signalosome (DLD-1, $P = 2.48 \times 10^{-6}$; HCT116, $P = 9.34 \times 10^{-6}$; HCC1954, $P = 4.54 \times 10^{-5}$; HMEC, $P = 3.2 \times 10^{-2}$), and the eukaryotic translation initiation factor 3 (eIF3) complex (DLD-1, $P = 1.42 \times 10^{-5}$; HCT116, $P = 7.98 \times 10^{-8}$; HCC1954, $P = 2.4 \times 10^{-4}$; HMEC, $P = 8.6 \times 10^{-3}$) were identified (Fig. 3B). Several key proteins in the ubiquitination and sumoylation pathways, including most of the cullins, were also identified. Multiple shRNAs against the same gene scored in the screen, which suggests that their effects are unlikely due to off-target effects.

We next validated *EIF3S10* and *RBX1*: two genes that are essential for viability in all four cell lines. For each gene, we included shRNAs that scored in the screen as well as additional shRNA sequences present in our library (table S2). Cells were infected with individual retroviral shRNAs, and cell viabilities were assessed (Fig. 3C). For each gene, all of the shRNAs that scored in the screen and many additional shRNAs gave antiproliferative phenotypes. Furthermore, the antiproliferative activity of the shRNAs correlated very well with the extent of target gene knockdown, as shown for

RBX1 (fig. S2A). Thus, these phenotypes are likely due to target gene knockdown rather than to off-target effects. This finding is consistent with a previous transfection-based screen with this library showing ~90% "on-target" efficiency (14).

In addition to the common set of shRNAs that impairs viability in all cell lines, we observed substantial numbers of genes that are selectively required for proliferation of each cell line (tables S3 to S6). These are particularly interesting because they may reflect differences in the underlying oncogenic context and therefore represent potential cancer-selective drug targets. We validated the gene *PPP1R12A*, which encodes a regulatory subunit of protein phosphatase 1 (PP1), for its selective requirement in HCC1954 but not DLD-1 cells (Fig. 4A). The PPP1R12A shRNA that gave the greatest depletion (shRNA 3) showed the strongest effect on HCC1954 cells but only marginally affected DLD-1 viability (Fig. 4B and fig. S2B). This finding was corroborated with four additional PPP1R12A small interfering RNAs (siRNAs). These shRNAs and siRNAs resulted in comparable knockdown of PPP1R12A protein in both cell lines (fig. S2B), indicating that the selective requirement for *PPP1R12A* by HCC1954 cells is not due to different degrees of protein knockdown. PPP1R12A has been shown to target PP1 isoforms to several substrates including myosin and merlin (15, 16).

Fig. 3. Genes commonly required for proliferation or survival of normal and cancer cells. Error bars represent SDs across triplicates. (A) Representative candidate shRNAs that reduce viability of all four cell lines. Multiple entries for the same gene indicate that multiple independent shRNAs have scored in the screen for that gene. For each shRNA, the normalized log$_2$ Cy3/Cy5 ratio (i.e., its relative abundance in the end sample versus initial sample) is given. (B) Examples of core cellular modules required in all four cell lines. Shown are the APC/C (top left), the COP9 signalosome (top right), the eIF3 translation initiation complex (bottom left), and E3 ubiquitin ligases (bottom right). Color scheme indicates genes required for viability of one (yellow), two (orange), three (red), or all four (black) cell lines. (C) Validation of selected shRNAs from the screen that reduce viability of all four cell lines. Candidate shRNAs were individually packaged into viruses and infected into cells in 96-well plates in independent triplicates. Cell viability was measured at day 9 after infection. All viability reductions were significant ($P < 0.05$), except where indicated by the number sign. FF, negative control shRNA targeting firefly luciferase.

Thus, PP1 activity reduction by PPP1R12A knockdown may lead to increased phosphorylation of key proteins that disrupt the viability of HCC1954 cells. Conversely, *PRPS2*, which encodes phosphoribosyl pyrophosphate synthetase 2 (an enzyme involved in nucleoside metabolism), is more selectively required by DLD-1 than HCC1954 cells (Fig. 4C and fig. S2C). These results suggest that distinct, genetic context–dependent vulnerabilities exist between these tumor cell lines.

Comparison between HCC1954 cells and normal HMECs also revealed a distinct subset of genes selectively required by each cell line (tables S4 and S6). Not surprisingly, a much larger set of 695 genes

is required by HMECs, likely reflecting the ability of normal cells to appropriately respond to various cellular stresses. Conversely, the relatively fewer genes that are required by the cancer cells underscore their ability to evade and overcome growth-inhibitory cues. Among the genes identified as essential for HMECs and HCT116 cells, but not DLD-1 or HCC1954 cells, is *HDM2*, which encodes the human homolog of MDM2 (the E3 ligase for p53) (Fig. 4D). HCC1954 and DLD-1 cells harbor inactivating mutations (Tyr^{163}→Cys^{163} and Ser^{241}→Phe^{241}, respectively) in the *TP53* gene and are therefore insensitive to MDM2 knockdown. Multiple MDM2 shRNAs selectively impaired the

viability of the p53 wild-type HMECs but not that of HCC1954 cells with mutant p53 (Fig. 4E and fig. S2D). Furthermore, we were able to pharmacologically validate this finding by interfering with MDM2 function using the inhibitor nutlin-3 (*17*) and recapitulating the sensitivity of these cells to MDM2 inactivation (Fig. 4F and fig. S2D).

Several genes appear to be selectively required by HCC1954 cells but not by HMECs (tables S4 and S6). Among these is the cell cycle regulator and spindle checkpoint kinase *BUB1* (Fig. 4G). We validated *BUB1* using both shRNAs and siRNAs to confirm that its knockdown is more detrimental to HCC1954 cells than to HMECs (Fig. 4H), despite similar levels of BUB1 protein reduction (fig. S2E). These results indicate that BUB1 is likely to play an integral role in supporting the oncogenic transformation of HCC1954 cells because they are more dependent on BUB1 function. One possible explanation for this enhanced dependency may be the near-tetraploid nature of the HCC1954 genome. As compared with the diploid HMECs, HCC1954 cells may rely more heavily on the spindle checkpoint to maintain genomic stability. Such a dependency is an example of "non-oncogene addiction" where cancer cells come to be highly dependent for growth and survival on the functions of genes that are themselves not oncogenes (*18*).

Our study and an accompanying paper (*19*) demonstrate that highly parallel dropout screens that use complex pools of shRNAs can be achieved with the use of HH barcodes in combination with highly penetrant vectors. Our ability to identify antiproliferative shRNAs specific to particular cell lines indicates that different cancer cells have distinct growth and survival requirements that cluster with cancer type. Targeting such key vulnerabilities is an attractive approach for cancer-selective therapeutics. The functional genetic approach demonstrated here presents an alternative and complementary effort to sequencing-based approaches such as the Cancer Genome Atlas and similar efforts, which focus on physical alterations of the cancer genome.

The most complex pool that we used contains 42,000 distinct shRNAs (fig. S3): an 80-fold increase in complexity as compared with that of previous dropout screens based on our designs (*6*). It is now conceivable for researchers to screen the entire human genome with ~3 shRNAs per gene using a pool of ~100,000 shRNAs in ~100 million cells. Thus, a large number of cancer and normal cell lines can be rapidly screened in this manner, through what we hope will become a "Genetic Cancer Genome Project," with the goal of generating cancer lethality signatures for different cancer types and thus identifying cancer type–specific lethal genes representing potential drug targets.

Fig. 4. Genes selectively required for proliferation or survival of cancer cells. Error bars represent SDs across triplicates. (**A**) Identification of *PPP1R12A* (one shRNA) and *PRPS2* (two shRNAs) as two genes that are selectively required by HCC1954 or DLD-1 cells, respectively, in the screen. (**B** and **C**) Validation of *PPP1R12A* (B) and *PRPS2* (C) as selectively required for viability of HCC1954 or DLD-1 cells, respectively. Cells were either infected with individual retroviral shRNAs or transfected with individual siRNAs in triplicates. Cell viability was measured at 9 days after infection (shRNA) or 4 days after transfection (siRNA) (*, $P < 0.05$). Luc, negative control siRNA targeting luciferase. PLK1, positive control siRNA targeting polo-like kinase 1. (**D**) Normalized log$_2$ ratios of an MDM2 shRNA in the screen. (**E**) shRNA knockdown of MDM2 selectively impairs the viability of HMECs. Cell viability was measured 9 days after infection with retroviruses expressing five different MDM2 shRNAs (*, $P < 0.05$). (**F**) Differential sensitivity of the four cell lines to the MDM2 inhibitor nutlin-3. Cell viability is reflective of their p53 status (HMECs and HCT116 cells, p53 wild-type; HCC1954 and DLD-1 cells, p53 mutant). Cell viability was measured after 4 days of nutlin-3 treatment (*, $P < 0.05$). ctrl, control. (**G**) Normalized log$_2$ ratios of a BUB1 shRNA from the screen. (**H**) Enhanced sensitivity of HCC1954 cells to BUB1 knockdown. Both shRNA (left) and siRNA (right) knockdown of BUB1 reduce HCC1954 cell viability but have no effect on HMEC viability. Cell viability was measured 4 or 9 days after transfection or infection, respectively (*, $P < 0.05$).

References and Notes
1. P. J. Paddison *et al.*, *Nature* **428**, 427 (2004).
2. J. M. Silva *et al.*, *Nat. Genet.* **37**, 1281 (2005).
3. T. F. Westbrook *et al.*, *Cell* **121**, 837 (2005).
4. N. Popov *et al.*, *Nat. Cell Biol.* **9**, 765 (2007).
5. I. G. Kolfschoten *et al.*, *Cell* **121**, 849 (2005).
6. V. N. Ngo *et al.*, *Nature* **441**, 106 (2006).
7. Materials and methods are available as supporting material on *Science* Online.
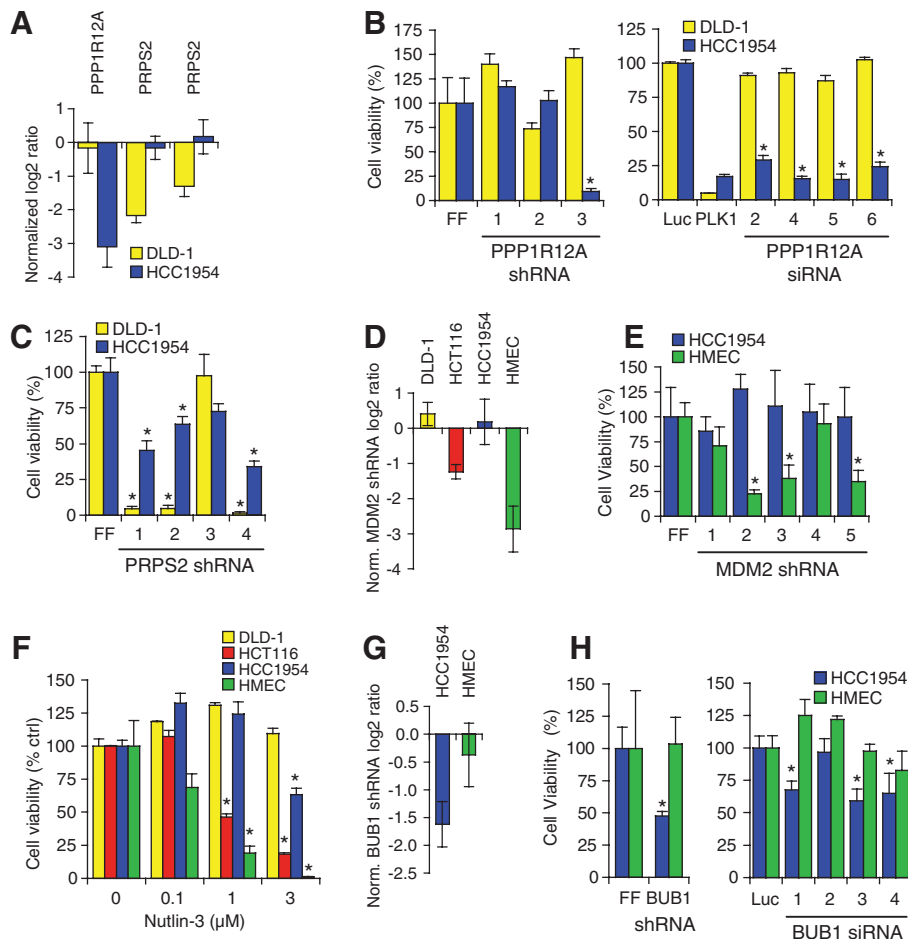
8. K. Berns et al., Nature **428**, 431 (2004).
9. T. R. Brummelkamp et al., Nat. Chem. Biol. **2**, 202 (2006).
10. T. Sjoblom et al., Science **314**, 268 (2006).
11. L. D. Wood et al., Science **318**, 1108 (2007).
12. R. A. Dickins et al., Nat. Genet. **37**, 1289 (2005).
13. G. K. Smyth, T. Speed, Methods **31**, 265 (2003).
14. V. M. Draviam et al., Nat. Cell Biol. **9**, 556 (2007).
15. M. Ito, T. Nakano, F. Erdodi, D. J. Hartshorne, Mol. Cell. Biochem. **259**, 197 (2004).
16. H. Jin, T. Sperka, P. Herrlich, H. Morrison, Nature **442**, 576 (2006).
17. L. T. Vassilev et al., Science **303**, 844 (2004).
18. N. L. Solimini, J. Luo, S. J. Elledge, Cell **130**, 986 (2007).
19. J. M. Silva et al., Science **319**, 617 (2008).

Cathepsin K–Dependent Toll-Like Receptor 9 Signaling Revealed in Experimental Arthritis

Masataka Asagiri,[1,2] Toshitake Hirai,[1,4] Toshihiro Kunigami,[1,4] Shunya Kamano,[1,5] Hans-Jürgen Gober,[1] Kazuo Okamoto,[1] Keizo Nishikawa,[1] Eicke Latz,[6] Douglas T. Golenbock,[6] Kazuhiro Aoki,[3] Keiichi Ohya,[3] Yuuki Imai,[7,9] Yasuyuki Morishita,[8] Kohei Miyazono,[8] Shigeaki Kato,[7,9] Paul Saftig,[10] Hiroshi Takayanagi[1,2]*

Cathepsin K was originally identified as an osteoclast-specific lysosomal protease, the inhibitor of which has been considered might have therapeutic potential. We show that inhibition of cathepsin K could potently suppress autoimmune inflammation of the joints as well as osteoclastic bone resorption in autoimmune arthritis. Furthermore, cathepsin K$^{-/-}$ mice were resistant to experimental autoimmune encephalomyelitis. Pharmacological inhibition or targeted disruption of cathepsin K resulted in defective Toll-like receptor 9 signaling in dendritic cells in response to unmethylated CpG DNA, which in turn led to attenuated induction of T helper 17 cells, without affecting the antigen-presenting ability of dendritic cells. These results suggest that cathepsin K plays an important role in the immune system and may serve as a valid therapeutic target in autoimmune diseases.

Both innate and adaptive immune systems contribute to the inflammation seen in autoimmune diseases, but the molecular mechanism underlying this process is not completely understood (1, 2). The cathepsins constitute a family of lysosomal cysteine proteases that were initially recognized as nonspecific scavengers of cellular proteins and that were also found to display cell type–specific functions (3, 4). Cathepsins L and S are fundamental in processing of major histocompatibility complex (MHC) class II antigens and MHC class II trafficking and maturation (3, 4). In contrast, cathepsin K is highly expressed in osteoclasts and is involved in degradation of bone matrices such as type I collagen (5). The loss-of-function mutation in the cathepsin K gene in humans causes pycnodysostosis, a rare genetic disorder characterized by impaired osteoclastic bone resorption (6). In mice, the targeted disruption of cathepsin K similarly results in the pycnodysostotic phenotype (4, 5, 7). Among matrix-degrading enzymes expressed in osteoclasts, cathepsin K is the only one for which an essential role in bone resorption has been clearly demonstrated in both mice and humans (8). Thus, cathepsin K remains a potential therapeutic target for the treatment of bone diseases such as osteoporosis and autoimmune arthritis, in which osteoclast activity is increased (9, 10).

Through screening, we obtained a potent orally active cathepsin K inhibitor named NC-2300 (Fig. 1A and figs. S1 to S3), which suppresses osteoclastic bone resorption both in vivo and in vitro (figs. S1 and S4). Computer-assisted simulation of the cathepsin K/NC-2300 complex indicated that NC-2300 blocks the active-site cleft where Cys^{25} and His^{162} of cathepsin K form the catalytic site (Fig. 1B and fig. S5). To test the effects of the inhibitor on disease models, we treated adjuvant-induced arthritis (AIA) in rats with oral administration of NC-2300 and compared the results with the effects of alendronate, which is one of the bisphosphonate compounds used clinically as an inhibitor of osteoclastic bone resorption. Bone loss in arthritis occurs mainly in two forms: bone erosion at the inflamed joints and periarticular osteoporosis (11). Radiological analysis revealed that NC-2300, but not alendronate, markedly suppressed bone erosion (Fig. 1C), although bone mineral density analysis showed that both compounds had a comparable inhibitory effect on periarticular osteoporosis (fig. S6A). NC-2300 also ameliorated paw swelling (Fig. 1D) and improved locomotive activity (fig. S6B) without affecting the onset rate of arthritis. NC-2300 reduced inflammation even when administered after the onset of disease (fig. S7). These results indicate that cathepsin K also functions in cells other than osteoclasts, allowing it to participate in autoimmune inflammation.

In AIA, local injection of adjuvant stimulates antigen presentation by dendritic cells (DCs), leading to T cell autoimmunity, the production of inflammatory cytokines by macrophages, and osteoclast-mediated bone destruction (9, 12, 13). The adjuvant effects are mainly dependent on the pathogen-associated molecular patterns (PAMPs)–induced activation of Toll-like receptor (TLR) signaling (14, 15). Therefore, we next analyzed the expression and function of cathepsin K in T cells, macrophages, and DCs. Cathepsin K mRNA was barely detected in nonadherent bone marrow (BM) cells or splenic T cells (Fig. 2A), and NC-2300 showed no effects on T cell activation (fig. S8A). Although macrophages have been reported to express cathepsin K (4), NC-2300 had no effects on the activation of BM-derived macrophages stimulated by PAMPs (fig. S8B). BM-derived DCs (BM-DCs) did express a detectable level of cathepsin K mRNA, although this was much lower than expression in osteoclasts (Fig. 2A). Nevertheless, cathepsin K activity was confirmed in DCs and was inhibited by NC-2300 (Fig. 2B).

To investigate whether cathepsin K has a role in antigen presentation in DCs, DCs were cultured with fluorescein isothiocyanate (FITC)–labeled ovalbumin. The uptake of ovalbumin-FITC was observed by flow cytometry in NC-2300–treated DCs as well as in nontreated cells (Fig. 2C). In addition, NC-2300–treated DCs stimulated proliferation of splenic T cells from ovalbumin-specific DO11.10 TCR transgenic mice to an extent similar to that of nontreated DCs (Fig. 2D). These results suggest that cathepsin K activity is not required for the antigen uptake, processing, or presentation by

[1]Department of Cell Signaling, Graduate School, Tokyo Medical and Dental University, Tokyo 113-8549, Japan. [2]Center of Excellence Program for Frontier Research on Molecular Destruction and Reconstruction of Tooth and Bone, Tokyo Medical and Dental University, Tokyo 113-8549, Japan. [3]Department of Hard Tissue Engineering, Section of Pharmacology, Graduate School, Tokyo Medical and Dental University, Tokyo 113-8549, Japan. [4]Nippon Chemiphar Co., Ltd., Saitama 341-0005, Japan. [5]Department of Orthopaedic Surgery, Juntendo University School of Medicine, Tokyo 113-8421, Japan. [6]Division of Infectious Diseases and Immunology, University of Massachusetts Medical School, Worcester, MA 01605, USA. [7]Institute of Molecular and Cellular Biosciences, University of Tokyo, Tokyo 113-0032, Japan. [8]Department of Molecular Pathology, Graduate School of Medicine, University of Tokyo, Tokyo 113-0033, Japan. [9]Exploratory Research for Advanced Technology, Japan Science and Technology Agency, Saitama 332-0012, Japan. [10]Biochemical Institute, Christian-Albrechts-University Kiel, D-24098 Kiel, Germany.

*To whom correspondence should be addressed. E-mail: taka.csi@tmd.ac.jp

DCs. Considering the crucial role of cathepsins L and S in antigen presentation (*3*, *4*, *16*), these results also indicate that the effects of NC-2300 on these cathepsins are negligible.

Complete Freund's adjuvant (CFA) is composed of killed mycobacteria, which contain various PAMPs including TLR2, TLR4, and TLR9 agonists (*14*, *17*). Therefore, we tested the responses of DCs to these PAMPs with or without cathepsin K inactivation. Enzyme-linked immunosorbent assay revealed that the production of cytokines such as interleukin-12 (IL-12) and IL-23 by DCs was significantly inhibited by NC-2300 when stimulated with oligodeoxynucleotides containing unmethylated CpG motif (CpG: the TLR9 ligand), but not with Malp2 and peptidoglycan (PGN) (the TLR2 ligands)

or lipopolysaccharide (LPS: the TLR4 ligand) (Fig. 2E). CpG-induced expression of IL-6, IL-12, and IL-23 in BM-DCs was down-regulated at the mRNA level (Fig. 2F), and IFN-β production by Flt3L-induced BM-DCs in response to CpG was suppressed by NC-2300 (Fig. 2F). These results suggest that cathepsin K plays an important role in the gene induction program regulated by TLR9 signaling.

Deoxyribonuclease treatment of the mycobacterial components in CFA greatly reduces the severity of AIA without affecting the induction efficiency, and this reduction is recovered by supplementation with CpG DNA (*18*, *19*), indicating that TLR9-mediated immune responses determine the severity of autoimmune inflammation. Because mycobacterial DNA can be detected weeks after CFA inoc-

ulation (*19*), CpG DNA may augment autoimmune inflammation throughout the course of arthritis. These observations lend support to the notion that cathepsin K–dependent TLR9 signaling contributes to autoimmune inflammation.

How does cathepsin K regulate TLR9 signaling? After being taken up into the cells, CpG DNA locates within the endosomal compartment, where CpG DNA binds to TLR9 (*14*, *18*). The binding of CpG with TLR9 results in the conformational change of TLR9 (*20*), leading to activation of MyD88 and downstream signaling such as the mitogen-activated protein kinase, interferon regulatory factor (IRF), and nuclear factor kappa B (NF-κB) pathways (*14*). Activation of these pathways leads to the production of inflammatory cytokines and up-regulation of cell-
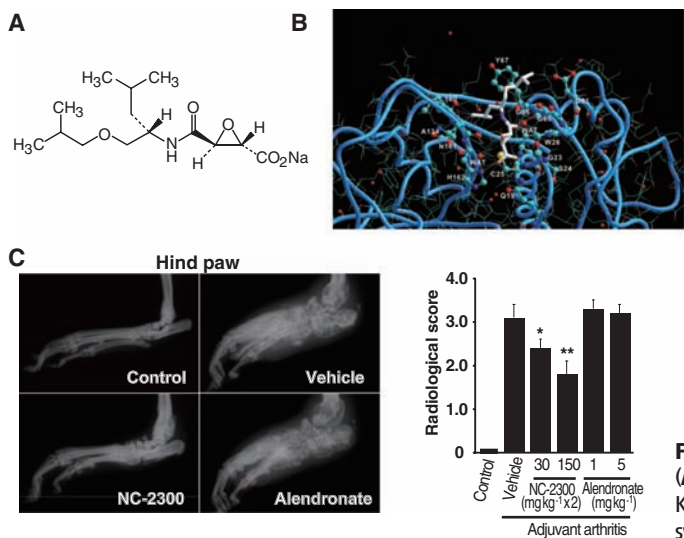

Fig. 1. Effect of the cathepsin K inhibitor NC-2300 on adjuvant-induced arthritis. (**A**) Structure of NC-2300. (**B**) Docking of NC-2300 toward the active site of cathepsin K. (**C**) Effect of NC-2300 on ankle-joint destruction. (**D**) Effect of NC-2300 on paw swelling. (C and D) $n = 10$, *$P < 0.05$; **$P < 0.01$; ***$P < 0.001$ (versus vehicle).
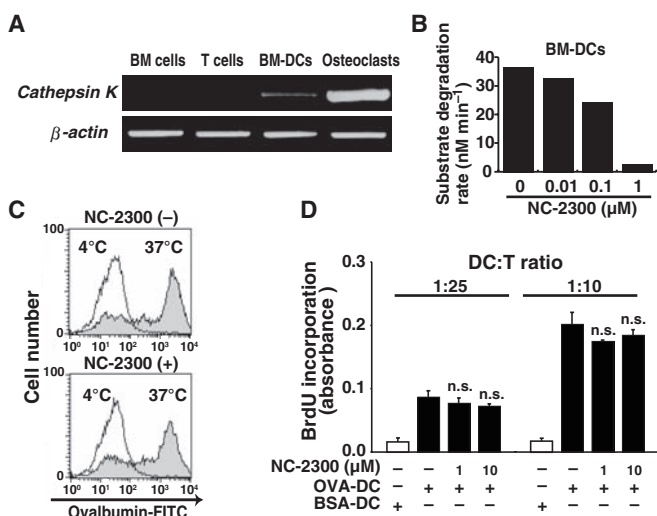

Fig. 2. Selective suppression of TLR9 response in dendritic cells (DCs) by cathepsin K inhibition. (**A**) Reverse transcription–polymerase chain reaction analysis of the cathepsin K mRNA. (**B**) Enzymatic activity of cathepsin K in BM-DCs. (**C**) Effect of NC-2300 on antigen uptake. (**D**) Effect of cathepsin K inhibition on the presentation of exogenous antigens by DCs. BrdU, bromodeoxyuridine; OVA, ovalbumin; BSA, bovine serum albumin; n.s., not significant. (**E**) Effects of cathepsin K inhibition on TLR-stimulated cytokine production. *$P < 0.05$; ***$P < 0.001$. (**F**) Effect of NC-2300 on the cytokine mRNA expression in CpG-treated BM-DCs.
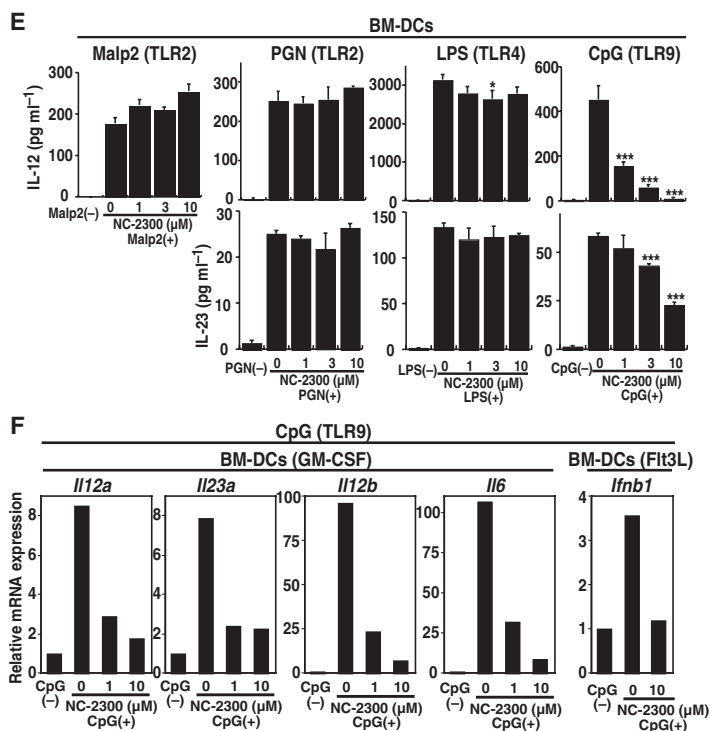
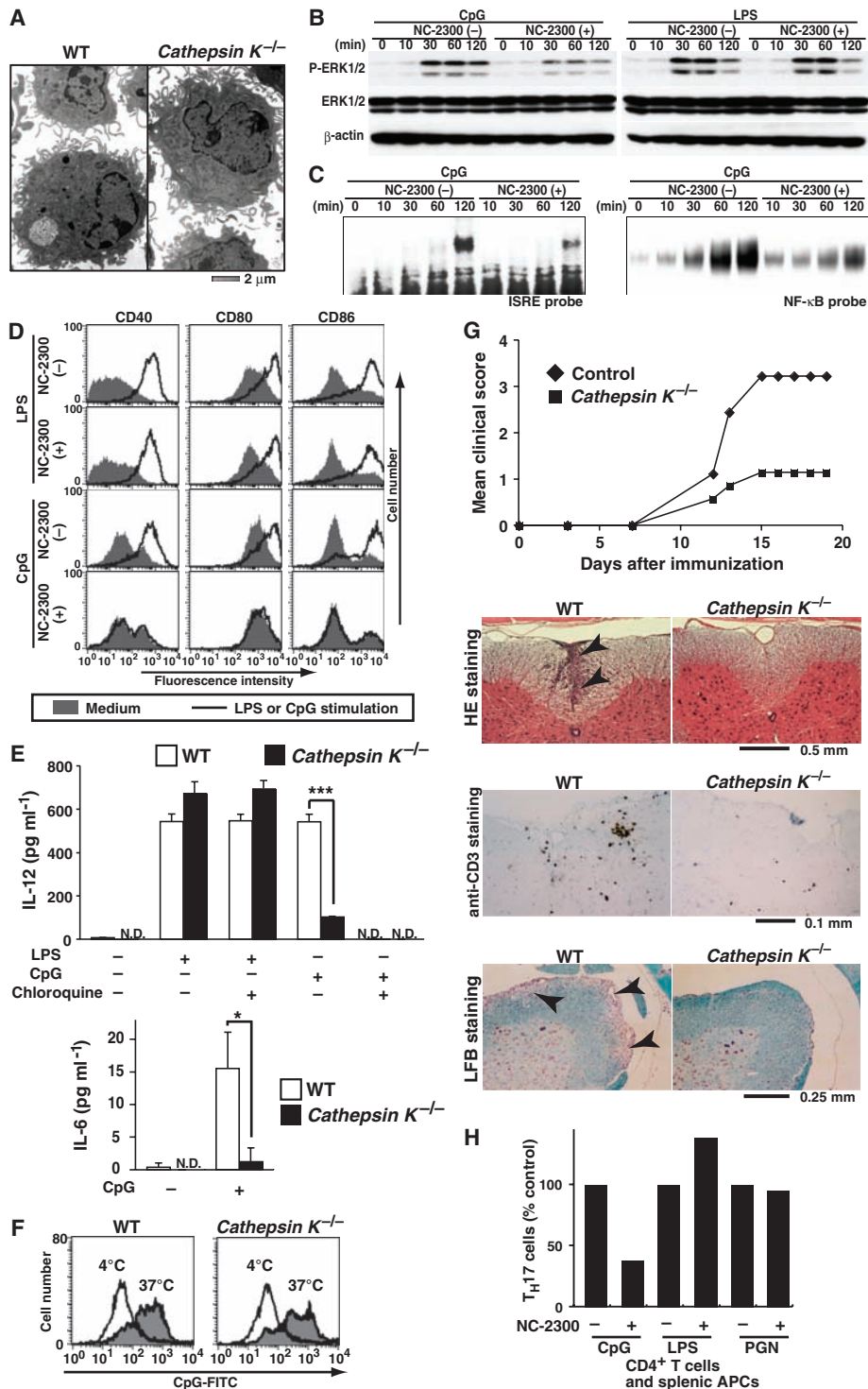**Fig. 3.** Cathepsin K regulation of TLR9-mediated immune responses and autoimmune inflammation. (**A**) Electron micrographs of WT and *cathepsin K⁻/⁻* BM-DCs. (**B**) Suppression of ERK phospho-rylation by NC-2300 in CpG-stimulated DCs. (**C**) Effect of NC-2300 on CpG-induced activation of transcription factors, IRF (left) and NF-κB (right) (electrophoretic mobility shift assay). ISRE, IFN-stimulated response element. (**D**) Impaired TLR9-mediated expression of cell-surface molecules in NC-2300–treated DCs. (**E**) Cytokine production in WT or *cathepsin K⁻/⁻* DCs in response to LPS or CpG. N.D., not detected. *$P < 0.05$; ***$P < 0.001$. (**F**) Uptake of CpG in *cathepsin K⁻/⁻* DCs. (**G**) Severity of EAE in *cathepsin K⁻/⁻* mice. Clinical scores (top). Lumbar spinal cord sections stained with hematoxylin and eosin (HE), immunostained for CD3, and stained with luxol fast blue (LFB) (bottom). Arrowheads indicate inflammatory cellular infiltrates (HE) and demyelinated areas (LFB). (**H**) Cathepsin K inactivation results in defective T$_H$17 polarization in response to CpG, but not to LPS or PGN.

surface molecules such as CD40, CD80, and CD86 in DCs (2, 21). Compounds that block endosomal acidification such as chloroquine inhibit CpG-driven signaling (14, 18). A normal number of BM-DCs were generated in BM cells derived from *cathepsin K⁻/⁻* mice (5, 22), and they exhibited a normal mor-phology (Fig. 3A), suggesting that cathepsin K is dis-pensable for the DC differentiation. CpG-induced, but not LPS-induced, phosphorylation of extracellular-signal regulated kinases 1 and 2 (ERK 1/2) in DCs was suppressed by cathepsin K inactivation (Fig. 3B). CpG-induced activation of IRF and NF-κB was also suppressed by cathepsin K inactivation in DCs (Fig. 3C). In addition, NC-2300 reduced ex-pression of CD40, CD80, and CD86 in DCs (Fig. 3D). CpG-induced, but not LPS-induced, production of cytokines such as IL-12 and IL-6 in *cathepsin K⁻/⁻* DCs was significantly suppressed (Fig. 3E). Thus, cathepsin K inactivation leads to the blockade of essentially all the downstream pathways of TLR9 signaling in DCs, suggesting that cathepsin K plays a critical role in the signaling events prox-imal to TLR9. However, the endocytosis of CpG was not affected in *cathepsin K⁻/⁻* DCs (Fig. 3F) and the endosomal acidification was not inhibited by NC-2300 (fig. S9), whereas chloroquine sup-pressed the cathepsin K activity in DCs (fig. S10).

It remains to be elucidated how cathepsin K regulates CpG-TLR9 signaling in the endosome, but the results of pharmacological inhibition indi-cate that the proteolytic activity of this enzyme is crucial for the mechanism. It is conceivable that cathepsin K may be involved in the degradation of proteins that inhibit the interaction between CpG and TLR9, or cathepsin K–mediated proteolysis may result in the conformational change of TLR9 that augments its signal transduction (20). Because fluorescence resonance energy transfer analysis showed that the CpG-induced conformational change of TLR9 was not affected in human embryonic kidney (HEK) 293 cells by NC-2300 (fig. S11), we cannot rule out the possibility that cathepsin K degrades a cytoplasmic protein that modifies the proximal TLR9 signaling. The observation that *cathepsin K⁻/⁻* DCs responded normally to TLR3 or TLR7/8 stimulation suggests that the role of cathepsin K is TLR9-specific in DCs (fig. S12).

To clearly demonstrate that cathepsin K plays a critical role in autoimmune inflammation in an osteoclast-independent manner, we subjected *ca-thepsin K⁻/⁻* mice to experimental autoimmune en-cephalomyelitis (EAE), in which TLR9 signaling plays an important role (23, 24). The frequency of the onset of EAE was not different between control and *cathepsin K⁻/⁻* mice, but the severity of the par-alytic symptoms was much lower in the *cathepsin K⁻/⁻* mice than in control mice (Fig. 3G, top). His-tological analysis of spinal cords demonstrated a marked decrease in inflammation (Fig. 3G, HE staining), T cell infiltration (Fig. 3G, anti-CD3 stain-ing), and demyelination (Fig. 3G, LFB staining) in *cathepsin K⁻/⁻* mice. Because T helper 17 (T$_H$17) cells play an essential role in the autoimmune inflam-mation in EAE (25), we examined the effect of ca-thepsin K inactivation on the ability of DCs to induce

T_H17 cells. The ability of DCs to induce T_H17 cells was markedly inhibited by cathepsin K inactivation when stimulated with CpG, but not with LPS or PGN (Fig. 3H). Taken together with the results on the role of cathepsin K in CpG-induced cytokine expression in DCs, the impaired induction of T_H17 cells by cathepsin K inactivation was caused, at least in part, by the reduced DC expression of cytokines that are involved in the induction and expansion of T_H17 cells such as IL-6 and IL-23 (25, 26).

Our results show that cathepsin K, which was thought to be an osteoclast-specific enzyme, plays a critical role in the immune system. Cathepsin K functions under the acidified conditions in the endosome, where engagement of CpG by TLR9 occurs, and plays an important role in the signaling events proximal to TLR9. Thus, careful attention should be paid to the side effects of cathepsin K inhibitors on the immune system in the treatment of osteoporosis, whereas they may have dual benefits in the treatment of autoimmune arthritis, the pathogenesis of which is dependent on both DCs and osteoclasts (9).

References and Notes
1. J. Banchereau, V. Pascual, A. K. Palucka, *Immunity* **20**, 539 (2004).
2. R. M. Steinman, D. Hawiger, M. C. Nussenzweig, *Annu. Rev. Immunol.* **21**, 685 (2003).
3. V. Turk, B. Turk, D. Turk, *EMBO J.* **20**, 4629 (2001).
4. K. Honey, A. Y. Rudensky, *Nat. Rev. Immunol.* **3**, 472 (2003).
5. P. Saftig et al., *Proc. Natl. Acad. Sci. U.S.A.* **95**, 13453 (1998).
6. B. D. Gelb, G. P. Shi, H. A. Chapman, R. J. Desnick, *Science* **273**, 1236 (1996).
7. W. Chen et al., *Hum. Mol. Genet.* **16**, 410 (2007).
8. M. Asagiri, H. Takayanagi, *Bone* **40**, 251 (2007).
9. H. Takayanagi, *Nat. Rev. Immunol.* **7**, 292 (2007).
10. Y. Yasuda, J. Kaleta, D. Brömme, *Adv. Drug Deliv. Rev.* **57**, 973 (2005).
11. S. Ochi et al., *Proc. Natl. Acad. Sci. U.S.A.* **104**, 11394 (2007).
12. R. Holmdahl et al., *Immunol. Rev.* **184**, 184 (2001).
13. A. Marshak-Rothstein, *Nat. Rev. Immunol.* **6**, 823 (2006).
14. S. Akira, S. Uematsu, O. Takeuchi, *Cell* **124**, 783 (2006).
15. A. Iwasaki, R. Medzhitov, *Nat. Immunol.* **5**, 987 (2004).
16. R. J. Riese et al., *J. Clin. Invest.* **101**, 2351 (1998).
17. S. B. Su et al., *J. Immunol.* **175**, 6303 (2005).
18. A. M. Krieg, *Annu. Rev. Immunol.* **20**, 709 (2002).
19. A. Ronaghy et al., *J. Immunol.* **168**, 51 (2002).
20. E. Latz et al., *Nat. Immunol.* **8**, 772 (2007).
21. I. Mellman, R. M. Steinman, *Cell* **106**, 255 (2001).
22. T. Nakamura et al., *Cell* **130**, 811 (2007).
23. M. Prinz et al., *J. Clin. Invest.* **116**, 456 (2006).
24. B. M. Segal, J. T. Chang, E. M. Shevach, *J. Immunol.* **164**, 5683 (2000).
25. H. Park et al., *Nat. Immunol.* **6**, 1133 (2005).
26. E. Bettelli et al., *Nature* **441**, 235 (2006).
27. We thank K. Takeda, M. Yamamoto, H. Oda, S. Imajoh-Ohmi, M. Matsuura, K. Nishioka, T. Kanda, K. Sato, S. Hida, Y. Iwai, H. Hemmi, H. Yanai, S. Kano, Y. Omatsu, T. Onodera, T. Honda, and A. Suematsu for discussion and assistance. This work was supported in part by a Grant-in-Aid for Creative Scientific Research from the Japan Society for the Promotion of Science (JSPS); grants for the Genome Network Project from the Ministry of Education, Culture, Sports, Science and Technology of Japan (MEXT); Grants-in-Aid for Scientific Research from JSPS and MEXT; Health Sciences Research Grants from the Ministry of Health, Labor and Welfare of Japan; and grants from the Deutsche Forschungsgemeinschaft and the NIH.

Systemic Leukocyte-Directed siRNA Delivery Revealing Cyclin D1 as an Anti-Inflammatory Target

Dan Peer,[1] Eun Jeong Park,[1] Yoshiyuki Morishita,[1] Christopher V. Carman,[2] Motomu Shimaoka[1]*

Cyclin D1 (CyD1) is a pivotal cell cycle–regulatory molecule and a well-studied therapeutic target for cancer. Although CyD1 is also strongly up-regulated at sites of inflammation, its exact roles in this context remain uncharacterized. To address this question, we developed a strategy for selectively silencing CyD1 in leukocytes in vivo. Targeted stabilized nanoparticles (tsNPs) were loaded with CyD1–small interfering RNA (siRNA). Antibodies to β_7 integrin (β_7 I) were then used to target specific leukocyte subsets involved in gut inflammation. Systemic application of β_7 I-tsNPs silenced CyD1 in leukocytes and reversed experimentally induced colitis in mice by suppressing leukocyte proliferation and T helper cell 1 cytokine expression. This study reveals CyD1 to be a potential anti-inflammatory target, and suggests that the application of similar modes of targeting by siRNA may be feasible in other therapeutic settings.

RNA interference (RNAi) has emerged as a powerful strategy for suppressing gene expression, offering the potential to dramatically accelerate in vivo drug target validation, as well as the promise to create novel therapeutic approaches if it can be effectively applied in vivo (1). Cyclin D1 (CyD1) is a key cell cycle–regulating molecule that governs the proliferation of normal and malignant cells (2, 3). In inflammatory bowel diseases, colon-expressed CyD1 is aberrantly up-regulated in both epithelial and immune cells (4, 5). Although CyD1 has also been implicated in promoting epithelial colorectal dysplasia and carcinogenesis, it is not clear whether leukocyte-expressed CyD1 contributes directly to the pathogenesis of inflammation and whether it might serve as a therapeutic target.

To address these questions, we used RNAi silencing of CyD1 in an experimental model of intestinal inflammation. A major limitation to the use of RNAi in vivo is the effective delivery of siRNAs to the target cells (6, 7). RNAi in leukocytes, a prime target for anti-inflammatory therapeutics, has remained particularly challenging, as these cells are difficult to transduce by conventional transfection methods and are often disseminated throughout the body, thus requiring systemic delivery approaches (8). One possibility is to use integrins, which are an important family of cell-surface adhesion molecules, as targets for siRNA delivery (8). Specifically, we have shown that antibody-protamine fusion proteins directed to the lymphocyte function-associated antigen–1 (LFA-1) integrin selectively delivered siRNAs to leukocytes, both in vitro and in vivo (8). However, whether an integrin-directed siRNA delivery approach can induce

[1]Immune Disease Institute and Department of Anesthesia, Harvard Medical School, 200 Longwood Avenue, Boston, MA 02115, USA. [2]Department of Medicine, Beth Israel Deaconess Medical Center, and Harvard Medical School, 330 Brookline Avenue, Boston, MA 02215, USA.

*To whom correspondence should be addressed. E-mail: shimaoka@cbrinstitute.org

Fig. 1. The processes involved in generating I-tsNPs. Multilamellar vesicle (MLV) [prepared as described in (9)] is extruded to form a unilamellar vesicle (ULV) with a diameter of ~100 nm. Hyaluronan is covalently attached
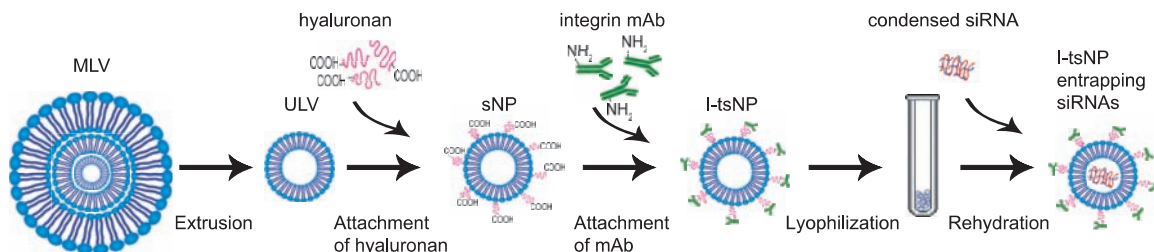
to DPPE in the ULV. A monoclonal antibody (mAb) to the integrin is covalently attached to hyaluronan, generating I-tsNP. siRNAs are entrapped by rehydrating lyophilized β_7 I-tsNP with water containing protamine-condensed siRNAs.

sufficiently robust silencing in vivo remains to be seen.

Building on the premise of integrin-targeted siRNA delivery, we developed liposome-based, β_7 integrin–targeted, stabilized nanoparticles (β_7 I-tsNPs) that entrap siRNAs (Fig. 1) (9). We began with nanometer-scale (~80 nm) liposomes, formed from neutral phospholipids to circumvent the potential toxicity common to cationic lipids and polymers used for systemic siRNA delivery (10). Hyaluronan was then attached to the outer surface of the liposomes, through covalent linkage to dipalmitoylphosphatidylethanolamine (DPPE), thereby stabilizing the particles both during subsequent siRNA entrapment (Fig. 1) and during systemic circulation in vivo (11). The resulting stabilized nanoparticles (sNPs) were successfully equipped with a targeting capacity by covalently attaching a monoclonal antibody against the integrins to hyaluronan (fig. S1). The antibody FIB504 (12) was selected to direct particles to β_7 integrins, which are highly expressed in gut mononuclear leukocytes (13).

We condensed siRNAs with protamine, a positively charged protein that has been used to enhance delivery of nucleic acids [e.g., DNA (14) and siRNA (15)]. β_7 I-tsNPs were loaded with siRNA cargo by rehydrating lyophilized particles in the presence of condensed siRNAs (9), thereby achieving ~80% entrapment efficacy while maintaining the nanodimensions of the particles (tables S1 and S2). β_7 I-tsNPs showed a measurable increase in their capacity to entrap siRNAs such that I-tsNPs carried ~4000 siRNA molecules per particle (~100 siRNA molecules per targeting antibody molecule) (table S1), as compared to an integrin-targeted single-chain antibody protamine fusion protein, which carried five siRNA molecules per fusion protein (8). The presence of hyaluronan was critical to maintaining the structural integrity of I-tsNPs during the cycle of lyophilization and rehydration (table S3 and fig. S2).

Cy3-siRNA encapsulated within β_7 I-tsNPs was efficiently bound and delivered to wild-type (WT) but not to β_7 integrin knockout (KO) splenocytes (Fig. 2A). Upon cell binding, β_7 I-tsNPs readily internalized and released Cy3-siRNA to the cytoplasm of both WT splenocytes (Fig. 2B) and the TK-1 lymphocyte cell line (fig. S3) but not that of β_7 integrin KO cells (Fig. 2B). Neither naked siRNA nor isotype control immunoglobulin G (IgG)–attached stabilized nanoparticles (IgG-sNPs) delivered Cy3-siRNA above background levels (Fig. 2, A and B). Using siRNA to Ku70, a ubiquitously expressed nuclear protein and reference target, we showed that Ku70-siRNA delivered by β_7 I-tsNPs induced potent gene silencing in splenocytes, whereas naked or IgG-sNP–formulated Ku70-siRNA did not (Fig. 2C) (additional results in figs. S4 to S7).

To investigate the ability of β_7 I-tsNPs to silence genes in vivo, we administered Ku70-siRNAs (2.5 mg per kilogram of body weight) entrapped in β_7 I-tsNPs by intravenous injection into mice and tested Ku70 expression in mononuclear leukocytes isolated from the gut and spleen after 72

hours (Fig. 2D) (9). Ku70-siRNAs delivered by β_7 I-tsNPs potently suppressed Ku70 expression in cells from the gut (including lamina propria and intraepithelial lymphocyte compartments) and spleen. No silencing was observed in cells from identically treated β_7 integrin KO mice, confirming the specificity to the β_7 integrin–expressing cells. Furthermore, naked siRNA as well as siRNA delivered in IgG-sNPs failed to induce detectable silencing in WT or KO mice.

We subsequently examined the biodistribution of ^{3}H-hexadecylcholesterol–labeled nanoparticles intravenously injected into healthy or diseased mice suffering from dextran sodium sulfate (DSS)–induced colitis (Fig. 2E) (9). IgG-sNPs showed very little distribution to the gut regardless of the presence of colitis. By contrast, a substantial portion (~10%) of β_7 I-tsNPs spread to the gut in healthy mice. The biodistribution of β_7 I-tsNPs to the gut selectively increased ~3.5-fold in the presence of colitis.

Using β_7 I-tsNPs, we next studied the effects of silencing by CyD1-siRNA (9). Treatment with β_7 I-tsNP–entrapped CyD1-siRNA reduced CyD1
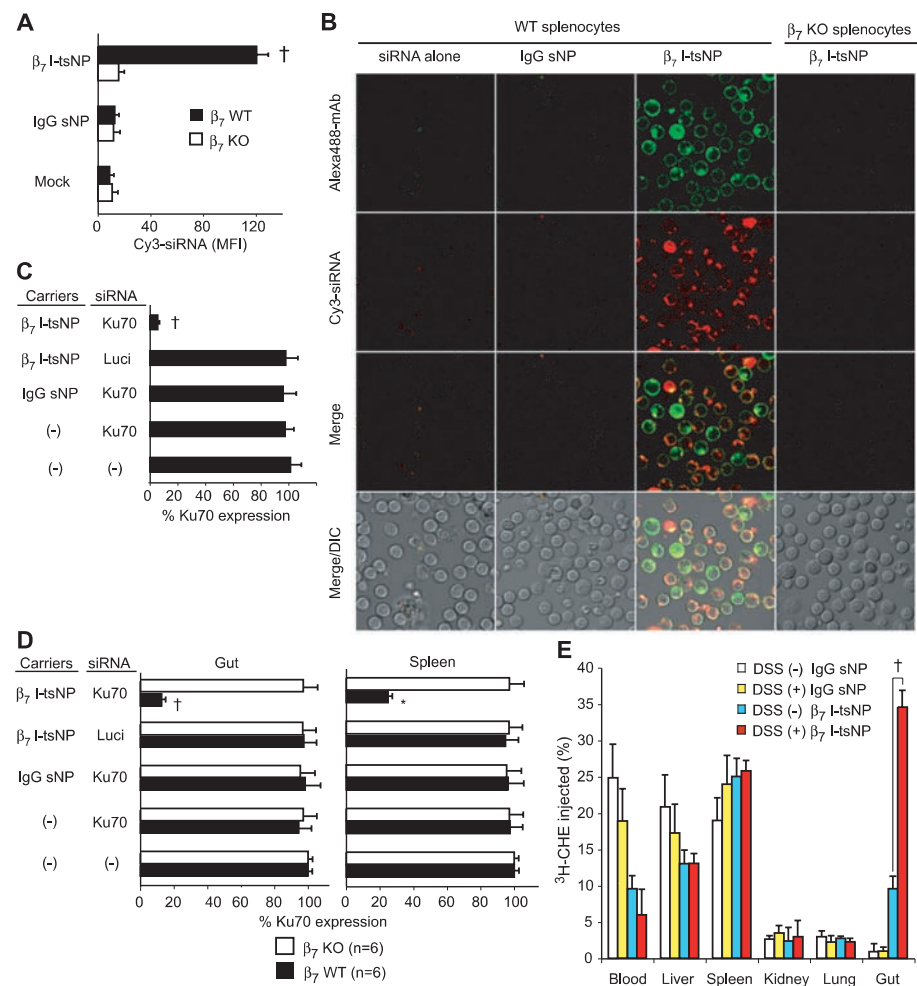

Fig. 2. β_7 I-tsNP delivers siRNAs in leukocytes in a β_7-specific manner. (**A**) Cy3-siRNA delivery via β_7 I-tsNP to WT, but not to β_7 knockout (KO), splenocytes as revealed by flow cytometry. (**B**) Confocal microscopy with differential interference contrast (DIC) morphologies showing the β_7 integrin–specific intracellular delivery of Cy3-siRNA. Images were acquired 4 hours after addition to splenocytes of naked Cy3-siRNA or Cy3-siRNA in Alexa 488–labeled β_7 I-tsNPs or IgG-sNPs. (**C**) Ku70-siRNA delivery with β_7 I-tsNP–induced silencing. Splenocytes were treated for 48 hours with 1000 pmol of Ku70-siRNAs or control luciferase (Luci)-siRNAs, delivered as indicated. (**D**) In vivo silencing of Ku70 in mononuclear cells from the gut and spleen of WT, but not KO, mice. siRNAs (2.5 mg/kg) entrapped as indicated were intravenously injected. Seventy-two hours after injection, Ku70 protein expression was determined by immunofluorescent cytometry after cell permeabilization and expressed as a percentage of Ku70 expression in mock-treated samples [(C) and (D)]. Data are expressed as the mean ± SEM of at least three independent experiments [(A), (C), and (D)]. *$P < 0.05$, †$P < 0.01$ versus mock-treated samples. (**E**) Biodistribution of ^{3}H-cholesterylhexadecylether (^{3}H-CHE)–labeled nanoparticles in mice with or without DSS-induced colitis. Pharmacokinetics and biodistribution were determined 12 hours after injection in a total of six mice per group in three independent experiments. Half-lives of β_7 I-tsNP in the blood of healthy and diseased mice were 4.3 and 1.8 hours, respectively. Preferential redistribution of β_7 I-tsNP to the inflamed gut is potentially advantageous for delivering siRNAs to treat intestinal inflammation. †$P < 0.01$.

mRNA expression in stimulated splenocytes, leading to potent suppression of proliferation (Fig. 3A). β_7 I-tsNPs–entrapped CyD1-siRNA (2.5 mg/kg) was then administered intravenously. Three days later, splenic and gut mononuclear leukocytes from mice treated with β_7 I-tsNP–entrapped CyD1-siRNA showed significantly decreased CyD1 mRNA and reduced proliferation (Fig. 3A).

Although CyD1-knockdown blocked agonist-enhanced expression of the T helper cell 1 (T_H1) cytokines interferon-γ (IFN-γ), interleukin-2 (IL-2), IL-12, and tumor necrosis factor–α (TNF-α), it did not alter the expression of the T_H2 cytokines IL-4 and IL-10 in either CD3/CD28- or phorbol 12-myristate 13-acetate (PMA)/ionomycin-stimulated splenocytes (Fig. 3B and fig. S8) or in PMA/ionomycin-stimulated TK-1 cells (fig. S9). The preferential inhibition of T_H1 cytokines was not observed with the CyD2- or CyD3-knockdown (fig. S9). To investigate whether the CyD1-knockdown suppressed T_H1 cytokine expression independently of its inhibitory effects on the cell cycle, we treated TK-1 cells with aphidicolin to arrest the cell cycle independently of the CyD1 status (Fig. 3C) (9). In aphidicolin-treated cells, PMA/ionomycin up-regulated CyD1 as well as T_H1 and T_H2 cytokines. CyD1-knockdown selectively suppressed T_H1 cytokine mRNA expression in aphidicolin-treated and PMA/ionomycin-activated cells (Fig. 3C). This cell cycle–independent suppression of T_H1 cytokines was also seen with the individual applications of four different CyD1-siRNAs that targeted non-

overlapping sequences in CyD1 mRNA (Fig. 3D), thereby ruling out the possibility that the blockade of T_H1 cytokines was due to an off-target effect. Thus, CyD1-knockdown could preferentially suppress pro-inflammatory T_H1 cytokine expression independently of any changes in the cell cycle.

We next studied CyD1-knockdown with β_7 I-tsNPs in vivo in DSS-induced colitis (9). Mice were intravenously injected with CyD1-siRNA (2.5 mg/kg) entrapped in β_7 I-tsNPs or IgG-sNPs at days 0, 2, 4, and 6. β_7 I-tsNP–delivered CyD1-siRNA potently reduced CyD1 mRNA to a level comparable with that of the uninflamed gut (Fig. 4D). CyD1-knockdown concomitantly suppressed mRNA expression of TNF-α and IL-12, but not of IL-10 (Fig. 4D). Notably, β_7 I-tsNP–delivered CyD1-siRNA led to a drastic reduction in intestinal tissue damage, to a potent suppression of leukocyte infiltration into the colon, and to a reversal in body weight loss and hematocrit reduction (Fig. 4, A to C, and fig. S10). The gut tissue of CyD1-siRNA/β_7 I-tsNP–treated animals exhibited normal numbers of mononuclear cells (Fig. 4C), suggesting that CyD1-knockdown does not induce pathologic cell death in the gut. CyD1-siRNAs entrapped in IgG-sNPs did not induce silencing in the gut, failing to alter cytokine expression in the gut or to reverse manifestations of colitis (Fig. 4, A to C) (additional results in figs. S11 and S12).

The anti-inflammatory effects of CyD1-knockdown in colitis are likely to be mediated both by suppressing the aberrant proliferation of

mucosal mononuclear leukocytes and by reducing the expression of TNF-α and IL-12, two pro-inflammatory T_H1 cytokines that are critical to the pathogenesis of colitis. The T_H2 cytokine IL-10 has been shown to suppress inflammation in colitis (16). Thus, its transformation from a relatively T_H1-dominant to a more T_H2-dominant phenotype appears to represent a critical and unexpected component of the potent colitis inhibition resulting from CyD1-knockdown. An important future goal will be to elucidate the molecular mechanisms underlying the effects of CyD1-knockdown, both on the induction of T_H2 polarization and on the specific type(s) of leukocytes responsible for reversal of colitis.

Entrapment of condensed siRNA inside these nanoparticles, in tandem with the targeting of the leukocyte β_7 integrin, which readily internalizes bound particles, enabled both highly efficient intracellular delivery and gene silencing in vivo. An effective in vivo siRNA dose of 2.5 mg/kg is one of the lowest doses reported to date for systemically targeted siRNA-delivery applications (17–22). Compared to other strategies, tsNPs offer the combined benefits of low off-target/toxicity with high cargo capacity (~4000 siRNA molecules per NP). Encapsulation of siRNA within the tsNPs seems to both protect siRNA from degradation (fig. S5) and prevent triggering of unwanted interferon responses (fig. S6). Antibodies coated on the outer surface of the NPs provided selective cellular targeting, while cell surface integrins proved to be effective anti-
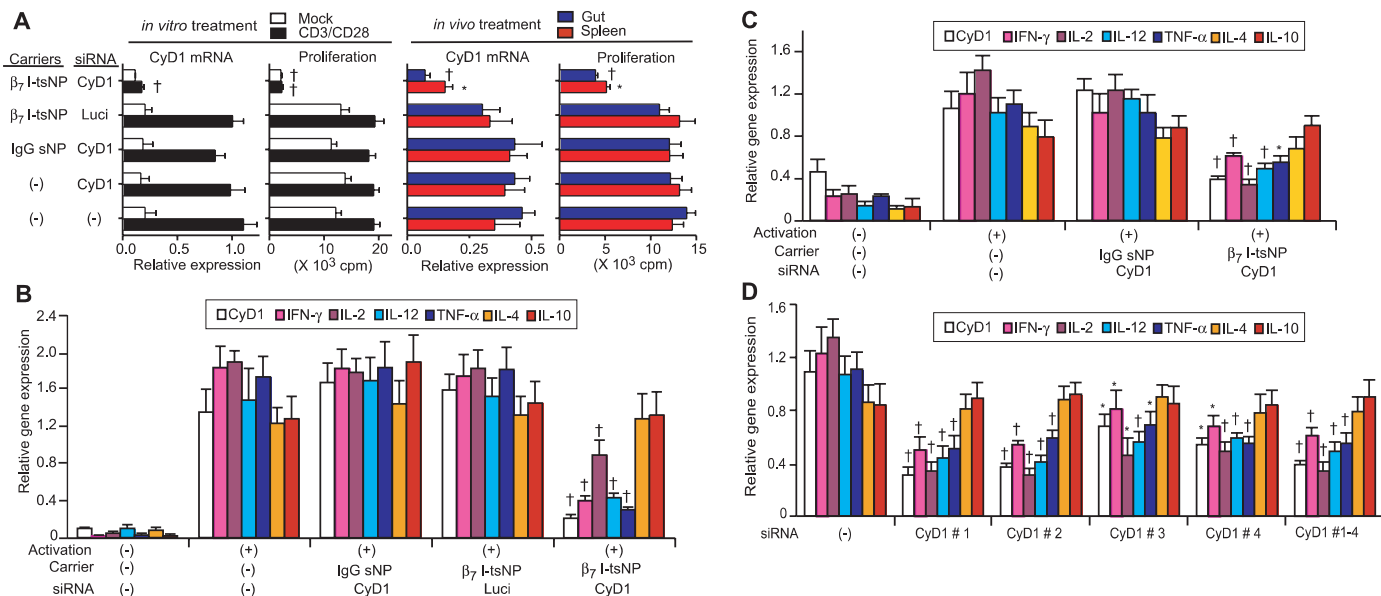

Fig. 3. Silencing of CyD1 by siRNA delivery with β_7 I-tsNPs and its effects on cytokine expression. (**A**) Silencing of CyD1 [measured by quantitative reverse transcription–polymerase chain reaction (qRT-PCR)] and its effects on proliferation (measured by [^{3}H]-thymidine incorporation). In in vitro treatments, splenocytes were examined after 72 hours' incubation with 1000 pmol of siRNAs delivered as indicated in the presence or absence of CD3/CD28 stimulation. In in vivo treatments, siRNAs (2.5 mg/kg) entrapped as indicated were intravenously injected into a total of six mice per group in three independent experiments. Seventy-two hours later, mononuclear cells harvested from the gut and spleen were examined. *$P < 0.05$, †$P < 0.01$ versus mock-treated samples. (**B**) CyD1-knockdown selectively suppresses T_H1 cytokine mRNA expression in splenocytes activated via

CD3/CD28. (**C**) CyD1-knockdown selectively suppresses T_H1 cytokine mRNA expression independently of its inhibitory effects on the cell cycle. In aphidicolin-treated TK-1 cells, in which the cell cycle was arrested, PMA/ionomycin–up-regulated T_H1 cytokine mRNA expression was selectively suppressed by CyD1-knockdown. (**D**) Cell cycle–independent suppression of T_H1 cytokines observed with individual applications of four different CyD1-siRNAs. (C) to (D) TK-1 cells were first treated for 12 hours with aphidicolin then with siRNAs (1000 pmol) delivered as indicated for another 12 hours in the presence of PMA/ionomycin and aphidicolin. (B) to (D) *$P < 0.05$, †$P < 0.01$ versus mock-treated activated cells. (A) to (D) mRNA levels for CyD1 and cytokines were measured by qRT-PCR. Data are expressed as the mean ± SEM of at least three independent experiments.
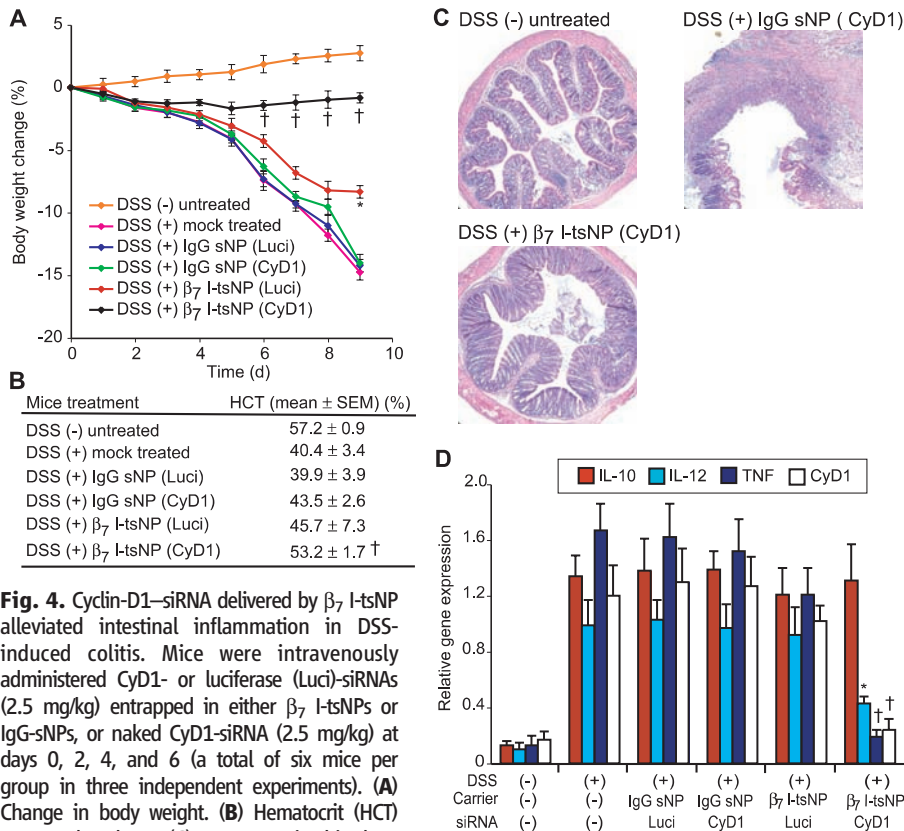
Fig. 4. Cyclin-D1–siRNA delivered by β_7 I-tsNP alleviated intestinal inflammation in DSS-induced colitis. Mice were intravenously administered CyD1- or luciferase (Luci)-siRNAs (2.5 mg/kg) entrapped in either β_7 I-tsNPs or IgG-sNPs, or naked CyD1-siRNA (2.5 mg/kg) at days 0, 2, 4, and 6 (a total of six mice per group in three independent experiments). (**A**) Change in body weight. (**B**) Hematocrit (HCT) measured at day 9. (**C**) Representative histology at day 9 (hematoxylin and eosin staining, magnification ×100). (**D**) mRNA expression of CyD1 and cytokines in the gut. mRNA expression was measured by qRT-PCR with homogenized colon samples harvested at day 9. (A), (B), and (D) Data are expressed as the mean ± SEM of three independent experiments. *$P < 0.05$, †$P < 0.01$ versus mock-treated mice with DSS-induced colitis.

body targets for both delivery and uptake of tsNPs. Thus, the I-tsNP approach may have broad applications not only for in vivo drug target validation, but also for potential therapies that are not limited to leukocytes or inflammatory settings.

References and Notes
1. E. Iorns, C. J. Lord, N. Turner, A. Ashworth, *Nat. Rev. Drug Discov.* **6**, 556 (2007).
2. D. W. Stacey, *Curr. Opin. Cell Biol.* **15**, 158 (2003).
3. M. Fu, C. Wang, Z. Li, T. Sakamaki, R. G. Pestell, *Endocrinology* **145**, 5439 (2004).
4. R. Yang, W. Bie, A. Haegebarth, A. L. Tyner, *Cell Cycle* **5**, 180 (2006).
5. H. van Dekken *et al.*, *Acta Histochem.* **109**, 266 (2007).
6. M. A. Behlke, *Mol. Ther.* **13**, 644 (2006).
7. D. M. Dykxhoorn, J. Lieberman, *Annu. Rev. Biomed. Eng.* **8**, 377 (2006).
8. D. Peer, P. Zhu, C. V. Carman, J. Lieberman, M. Shimaoka, *Proc. Natl. Acad. Sci. U.S.A.* **104**, 4095 (2007).
9. Materials and methods are available as supporting material on *Science* Online.
10. C. R. Dass, *J. Mol. Med.* **82**, 579 (2004).
11. D. Peer, A. Florentin, R. Margalit, *Biochim. Biophys. Acta* **1612**, 76 (2003).
12. D. P. Andrew *et al.*, *J. Immunol.* **153**, 3847 (1994).
13. S. K. Shaw, M. B. Brenner, *Semin. Immunol.* **7**, 335 (1995).
14. F. L. Sorgi, S. Bhattacharya, L. Huang, *Gene Ther.* **4**, 961 (1997).
15. S. D. Li, L. Huang, *Mol. Pharm.* **3**, 579 (2006).
16. J. O. Lindsay, A. Sandison, P. Cohen, F. M. Brennan, H. J. Hodgson, *Dig. Dis. Sci.* **49**, 1327 (2004).
17. J. Soutschek *et al.*, *Nature* **432**, 173 (2004).
18. D. V. Morrissey *et al.*, *Nat. Biotechnol.* **23**, 1002 (2005).
19. R. M. Schiffelers *et al.*, *Nucleic Acids Res.* **32**, e149 (2004).
20. E. Song *et al.*, *Nat. Biotechnol.* **23**, 709 (2005).
21. F. Takeshita *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **102**, 12177 (2005).
22. J. D. Heidel *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **104**, 5715 (2007).
23. We thank J. Lieberman for critically reading manuscript and discussion; R. Margalit, R. S. Langer, and P. Sicinski for discussions; and Y. Imai, A. Zur, P. Sage, and R. Yoo for technical assistance. D.P. is supported by the Dorot Foundation and Pfizer Inc. Y.M. is supported by the Uehara Memorial Foundation. This work was supported by the Arthritis Foundation (C.V.C.) and NIH grants HL048675 and AI63421 (M.S.).

Direct Observation of Hierarchical Folding in Single Riboswitch Aptamers

William J. Greenleaf,[1]* Kirsten L. Frieda,[2] Daniel A. N. Foster,[4] Michael T. Woodside,[4,5]*† Steven M. Block[1,3]†

Riboswitches regulate genes through structural changes in ligand-binding RNA aptamers. With the use of an optical-trapping assay based on in situ transcription by a molecule of RNA polymerase, single nascent RNAs containing *pbuE* adenine riboswitch aptamers were unfolded and refolded. Multiple folding states were characterized by means of both force-extension curves and folding trajectories under constant force by measuring the molecular contour length, kinetics, and energetics with and without adenine. Distinct folding steps correlated with the formation of key secondary or tertiary structures and with ligand binding. Adenine-induced stabilization of the weakest helix in the aptamer, the mechanical switch underlying regulatory action, was observed directly. These results provide an integrated view of hierarchical folding in an aptamer, demonstrating how complex folding can be resolved into constituent parts, and supply further insights into tertiary structure formation.

Riboswitches are elements of mRNA that regulate gene expression through ligand-induced changes in mRNA secondary or tertiary structure (*1*, *2*). This regulation is accomplished through the binding of a small metabolite to an aptamer in the 5′-untranslated region of the mRNA, which causes conformational changes that alter the expression of downstream genes.

Riboswitch-dependent regulatory processes depend crucially on the properties of aptamer folding; the kinetics and thermodynamics of folding are therefore of central importance for understanding function.

Among the simplest riboswitches are those regulating purine metabolism, which have aptamers with "tuning fork" structures (*3*, *4*) that bind ligands at a specific residue in a pocket formed by a three-helix junction. The junction is thought to be preorganized by numerous tertiary contacts, including interactions between two hairpin loops, but the binding pocket itself is likely stabilized only upon ligand binding (*4*–*10*). Ligand binding also stabilizes a nearby helix (*3*–*5*), sequestering residues that would otherwise participate in an alternate structure affecting gene expression (e.g., terminator or anti-terminator hairpins, ribosome binding sequences). Features such as ligand specificity (*6*, *11*) and its structural basis (*6*, *7*), the rates and energies for ligand binding and dissociation (*12*), the kinetics of loop-loop formation (*10*), and the interplay of structural preorganization and induced fit (*7*–*9*) have recently been investigated. These studies, how-

ever, focused on isolated steps in folding, typically using ligand analogs or investigating aptamers from different organisms. Here we obtain, from a single set of measurements, an integrated picture of secondary and tertiary structure formation, as well as ligand binding, in the aptamer of the *pbuE* adenine riboswitch from *Bacillus subtilis*, by observing folding and unfolding trajectories of individual molecules subjected to controlled loads in a high-resolution, dual-trap optical tweezers apparatus (*13*).

Single-molecule force spectroscopy, which measures the extension of a molecule as it unfolds and refolds under tension, furnishes a tool for probing structural transitions: Extension changes can be related to the number of nucleotides involved in folding. Furthermore, the effects of force on reaction equilibria and kinetics allow the shapes of the folding landscapes to be determined in detail (*14–16*). The complete folding process, starting from a fully unfolded state (not usually probed in conventional RNA folding studies), can also be observed. This initial configuration is especially relevant to riboswitches, because aptamers fold cotranscriptionally from an initially unstructured state. Because of the tight coupling between folding and transcription, the assay was designed to measure folding of mRNA transcribed in situ (*17*). A single *Escherichia coli* RNA polymerase (RNAP) molecule, transcriptionally stalled downstream of the promoter region on a DNA template (Fig. 1A), was attached to a bead held in one optical trap (Fig. 1B). The 29-nucleotide (nt) initial RNA transcript emerging from the RNAP was hybridized to the complementary cohesive end of a 3-kb double-stranded DNA (dsDNA) "handle" attached to a bead held in the other trap, creating a "dumbbell" geometry that allowed forces to be applied between the RNAP and the 5′ end of the RNA (*18*). Force-extension curves (FECs), showing the molecular extension measured as a function of force as the traps were moved apart at a constant rate, confirmed that this initial transcript was unstructured (Fig. 1C).

After constructing the dumbbells, transcription was restarted by introducing nucleoside triphosphates. The DNA template coded for the *pbuE* adenine riboswitch aptamer downstream of the initial transcript (Fig. 1A). Once the aptamer sequence was transcribed, RNAP was prevented from further elongation by a roadblock consisting of a streptavidin molecule bound to a 5′-terminal biotin label on the template (Fig. 1D). FECs measured immediately after aptamer transcription (Fig. 1E) revealed a characteristic series of sawtooth features that arise from contour length increases as specific structural

elements unfold (*19*). In the absence of adenine, two small unfolding events were typically observed (Fig. 1E, black). These features are produced by the unfolding of the two stable hairpins in the secondary structure, P3 and P2 (Fig. 1E, inset). The interactions that underpin tertiary structure by holding these hairpin loops together and structuring the binding pocket in the triple-helix junction are present only transiently in the absence of adenine (*5, 8, 10*). The contour length changes associated with these features, 17 ± 2 nt (P3) and 22 ± 2 nt (P2), are consistent with the values expected for these hairpins (19 and 21 nt, respectively) (*18*). In the presence of adenine, some FECs were identical to those observed in its absence, indicating in these cases that adenine was not bound to the aptamer. More commonly, however, larger unfolding distances at higher forces were observed, corresponding to adenine-induced stabilization of the folded structure. In the latter case, the aptamer usually unfolded cooperatively in a single event (Fig. 1E, blue), but sometimes through an intermediate state (Fig. 1E, red).

Unfolding from the fully folded state was analyzed in more detail by collecting multiple FECs from the same molecule. Overlaying 800 FECs shows that the aptamer unfolds over a wide distribution of forces (Fig. 2A), as expected for a nonequilibrium measurement (*20*). Three states were clearly seen: the folded and unfolded states, and an intermediate state. We fit the FECs with two worm-

like chains (WLCs) in series: one for the dsDNA handle (*21*) and the other for the single-stranded RNA (*22*), assuming a contour length of 0.59 nm/nt for RNA (*23*). When the aptamer unfolded fully, 62 ± 1 nt were released (*18*), in agreement with the 63-nt aptamer length. The intermediate state is 23 ± 1 nt shorter than the unfolded state, suggesting that it corresponds to a folded 21-nt P2 helix. The equilibrium free energy of the aptamer, computed by the method of Jarzynski (*24, 25*) from the nonequilibrium work done to unfold it (fig. S2), is 18 ± 2 kcal/mol (*18*). For comparison, the free energy predicted for the secondary structure in Fig. 1E is only ~12 ± 1 kcal/mol (*10*), indicating that tertiary contacts and ligand binding stabilize the aptamer by an additional ~6 ± 2 kcal/mol, in reasonable agreement with earlier measurements of the binding energy of 2-aminopurine (2AP), an adenine analog (*12*).

The distribution of forces, $p(F)$, for unfolding the fully folded aptamer (Fig. 2B) is well fit by an expression derived by Dudko *et al.* (*20*) for unfolding at fixed loading rate, parameterized by k_{off}, the unfolding rate at $F = 0$; $\Delta x^{\ddagger}$, the distance to the transition state from the folded state; and $\Delta G^{\ddagger}$, the height of the energy barrier (*18*). More than 3000 FECs measured for eight molecules, at loading rates varying from ~10 to 200 pN/s, yielded an unfolding rate $k_{off} \sim 0.04$ s^{-1} (ln $k_{off} = -3.5 \pm 1$), similar to the value of 0.15 s^{-1} measured previously by bulk kinetic methods (*12*). The activation
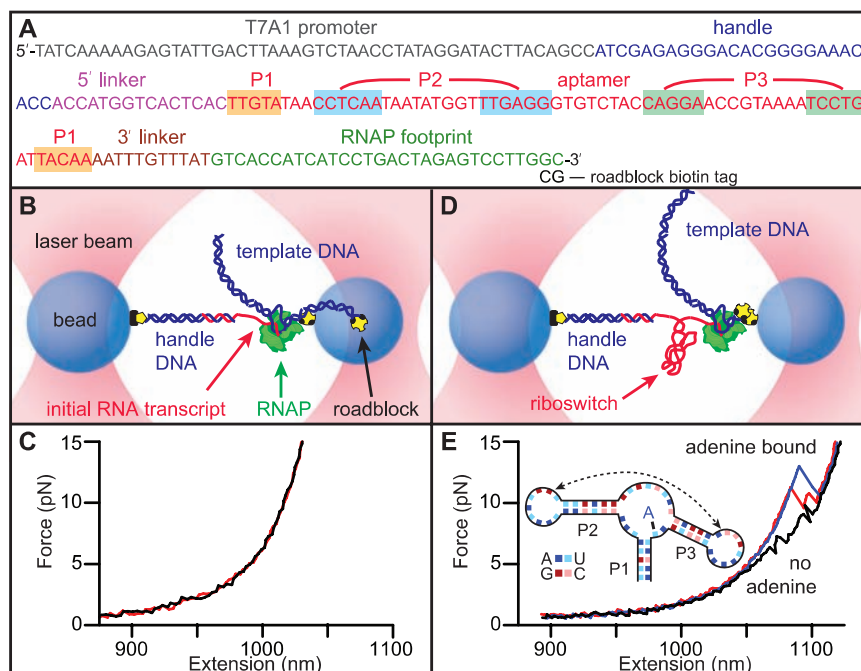

Fig. 1. (**A**) DNA template used for RNA transcription, showing the sequence of the nontranscribed promoter, the 25-bp section hybridizing with the DNA handle, the *pbuE* riboswitch aptamer (base-paired helices highlighted) flanked by short linkers, and the footprint of RNAP when stalled by the terminal roadblock. (**B**) Schematic of the optical trapping assay showing experimental geometry, with stalled RNAP and initial RNA transcript hybridized to the dsDNA handle (not to scale). (**C**) Two FECs obtained before aptamer transcription show little or no structure in the initial transcript. (**D**) Template DNA is transcribed in situ, producing an aptamer transcript, after which RNAP is stalled by a streptavidin molecule bound to the biotin-based roadblock. (**E**) FECs obtained after transcription show unfolding transitions in the aptamer. Without adenine, two events are seen (black), corresponding to the unfolding of hairpins P2 and P3 (inset). With adenine bound to the aptamer, larger unfolding events are observed (blue), sometimes involving an intermediate state (red).

[1]Department of Applied Physics, Stanford University, Stanford, CA 94305, USA. [2]Biophysics Program, Stanford University, Stanford, CA 94305, USA. [3]Department of Biological Sciences, Stanford University, Stanford, CA 94305, USA. [4]Department of Physics, University of Alberta, Edmonton AB, T6G 2G7, Canada. [5]National Institute for Nanotechnology, National Research Council of Canada, Edmonton, AB, T6G 2M9, Canada.

*These authors contributed equally to this work.
†To whom correspondence should be addressed. E-mail: sblock@stanford.edu (S.M.B); michael.woodside@nrc.ca (M.T.W.)

energy, $\Delta G^{\ddagger}$, was 17 ± 4 kcal/mol, in agreement with a previous result for the unbinding of 2AP (*12*). The distance to the transition state $\Delta x^{\ddagger}$ was 2.1 ± 0.2 nm. Given an extension of ~0.42 nm/nt at the average unfolding force of ~15 pN, this result indicates that the transition state involves the unzipping of ~2.5 base pairs (bp) in helix P1, suggesting that the G:C base pair in P1 (Fig. 1E, inset) represents a structural keystone: Both the binding pocket and triple-helix junction unfold once it is disrupted. Isolated G:C base pairs located 3 to 4 bp from the loop of P1 are found in the other purine

riboswitches, suggesting that they may be an important structural feature of this class of aptamers.

The kinetics of refolding and ligand binding were probed by observing the fraction of FECs showing the unfolding signature of the fully folded, adenine-bound aptamer, as a function of adenine concentration and the variable time interval during which refolding could occur between successive measurements (Fig. 3). We fit these data to a minimal, two-step model (Fig. 3, inset): formation of an intermediate structure competent to bind adenine (taken to be effectively irreversible) followed by

adenine binding. The complete folding process involves a hierarchy of several steps, including folding of the three helices, formation of the loop-loop contacts and the adenine binding pocket, and binding of adenine. At $F = 0$, however, helix formation should be fast compared to formation of the tertiary
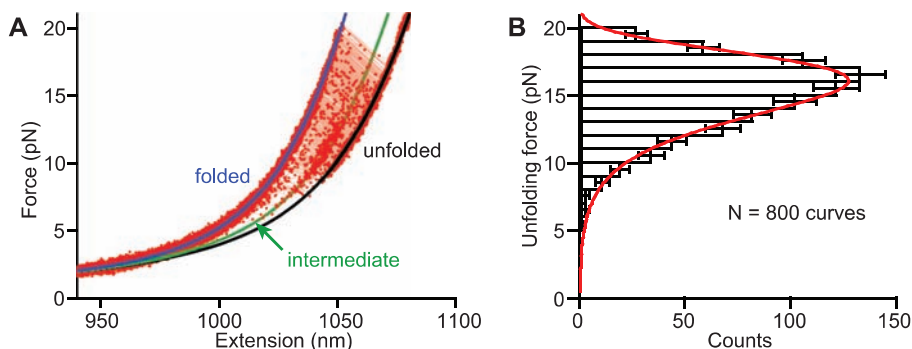

Fig. 3. Kinetics of aptamer refolding and binding. The fraction of FECs corresponding to the fully folded, adenine-bound aptamer (identified by the appropriate unfolding signature) for various adenine concentrations as a function of the variable time delay for refolding between pulls. Solid curves display the global fit to a minimal three-state kinetic scheme (**inset**): U, unfolded; A-comp, competent to bind adenine; F, folded (adenine-bound aptamer).

Fig. 2. (**A**) Nonequilibrium FECs for folded aptamer display a wide distribution of unfolding forces. WLC fit to the folded state (blue), and double WLC fits to the intermediate (green) and unfolded (black) states, indicate contour length changes of 39 ± 1 nt and 62 ± 1 nt for unfolding to the intermediate and unfolded states, respectively. (**B**) The unfolding force distribution is fit by a model returning the unfolding rate, along with the location and height of the energy barrier to unfolding.
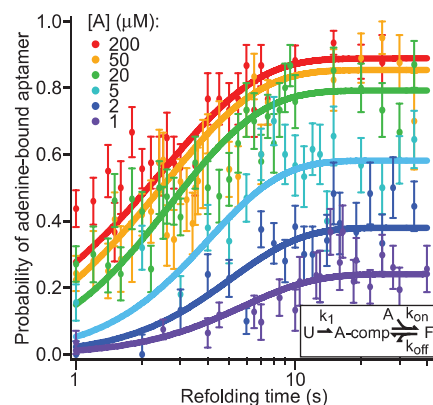
Fig. 4. Aptamer states and energetics determined by refolding at constant force. (**A**) As force is reduced, first P2 refolds (red), then P3 folds (orange). At lower forces, P2 and P3 interact to form a binding pocket and adenine binds, generating two additional states (green). The adenine-bound state is stable over many seconds, even at 5 pN load (blue). (**B**) Histograms of complete trajectories at different forces, with extension changes scaled by the force-dependent extension per nucleotide. Dashed lines indicate distinct states; the A-comp state is rarely populated. (**C** and **D**) Refolding trajectory and histograms in the absence of adenine. P2 and P3 folding occur as with adenine, but the A-comp state is highly populated at low force, whereas the folded state is very unstable, even at low force (purple). (**E**)

Quantitative energy landscapes for aptamer folding at 6.5 pN, reconstructed from the experimental data in the presence (red) and absence (black) of adenine. The five potential wells correspond to five observed folding states, illustrated by cartoons. Adenine binding only appreciably affects the barrier and energy of the folded state.
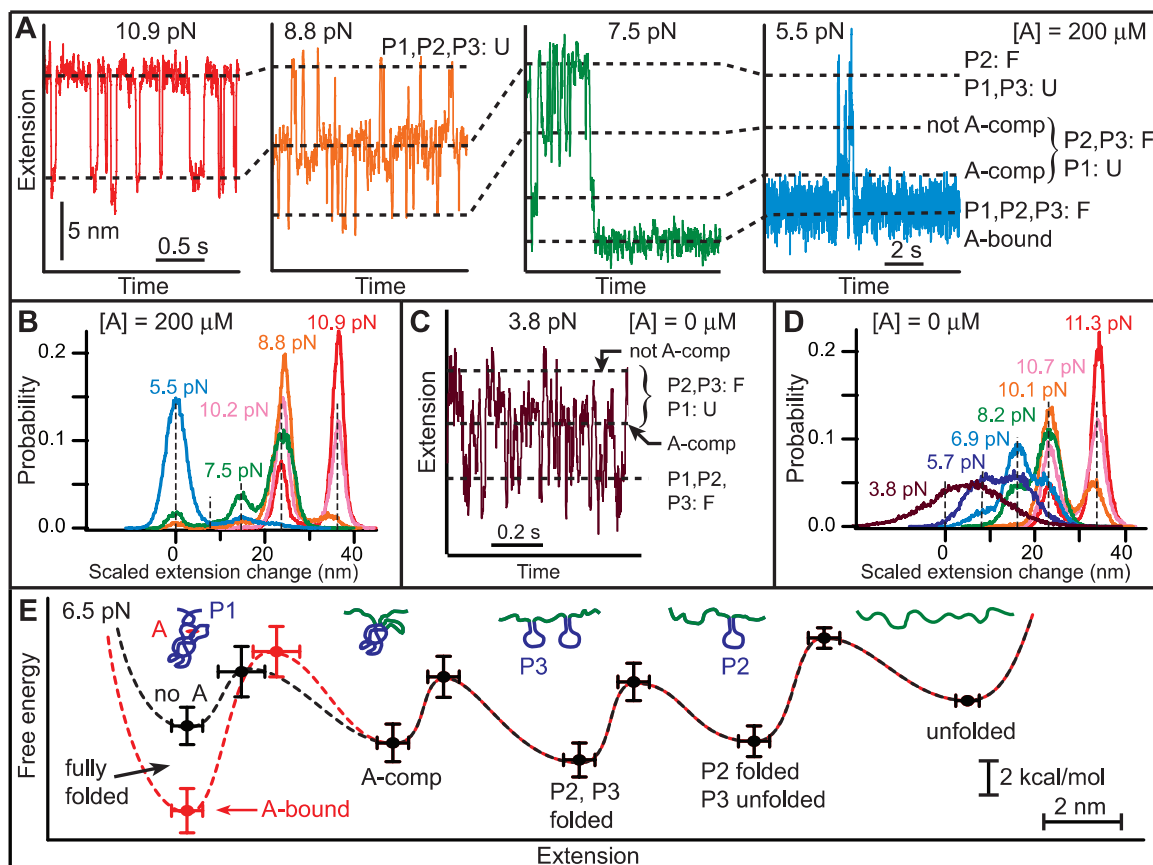
interactions creating the binding pocket; hence, we model this process using just three distinct states. A global fit of data to the time-dependent folding probabilities returned $k_1 = 0.4 \pm 0.05$ s^{-1}, $k_{off} = 0.2 \pm 0.05$ s^{-1}, and $k_{on} = 8 \pm 1 \times 10^4$ M^{-1} s^{-1}. The value of k_{off} is similar to that obtained above, and k_{on} is close to the value measured by bulk experiments (12). The slow folding rate implies that both aptamer folding and adenine binding occur on the same time scale as transcription itself, supporting the hypothesis that the function of this riboswitch is governed by folding and binding kinetics rather than equilibrium thermodynamics (10, 26).

The multiple steps in the overall folding reaction were studied in greater detail by unfolding the aptamer completely and monitoring its extension under constant force with a passive force clamp (27) as the force was reduced stepwise every 1 to 2 min. Observing the transitions in the refolding process individually, based on their different energies and time scales, four distinct steps were seen (Fig. 4). The first folding event, at ~9 to 11 pN (Fig. 4A, red), involves length changes and force-dependent kinetics consistent with folding the 21-nt hairpin P2, as predicted by an energy landscape model for hairpin folding (16, 18). The second folding step, at ~7 to 8 pN (Fig. 4A, orange), matches the properties expected for folding the 19-nt hairpin P3. The identification of these steps with the folding of P2 and P3 was confirmed by blocking the folding of each hairpin separately with antisense oligomers (fig. S3). We speculate that P2, the first fully transcribed element, is also the most stable in order to ensure that it can form in the presence of competing, alternative secondary structures in the upstream mRNA that might delay or prevent the formation of the proper aptamer structure.

In contrast to the adenine-independent events described above, the two folding transitions observed at lower forces were found to be adenine-dependent. For forces below ~7 pN at saturating adenine concentrations, the aptamer spent considerable time in the shortest-extension state (Fig. 4A, green and blue), which we identify as the folded, adenine-bound state. That identification was confirmed by measuring a contour length change of 63 ± 1 nt when the 63-nt aptamer was completely unfolded from this state. The contour length change between this state and the one with only P2 and P3 folded, 21 ± 2 nt, is consistent with the 23 nt that are not involved in P2 and P3 folding. In addition, a transient intermediate was observed between these two states, 14 ± 1 nt from the folded state. Because there are 15 nt in and adjacent to P1 (Fig. 1A), we identify this intermediate as a state where P2 and P3 are folded and the adenine binding pocket is preorganized by tertiary contacts, sequestering the nucleotides between P2 and P3, but P1 remains unfolded. The extensions of all five states (fully unfolded, P2 folded, P2/P3 folded, P1 unfolded, and fully folded), scaled by the fractional extension per nucleotide at a given force, are evident in histograms of records (Fig. 4B).

Constant-force–extension records in the absence of adenine (Fig. 4, C and D) indicate very different behavior at low forces: The P1-unfolded state is strongly populated below ~6 pN, whereas the folded state is only well populated below ~4 pN. Even at such low forces, the folded-state lifetime is short, with a rapid equilibrium between folded and P1-unfolded states. These differences can be understood if the P1-unfolded state is the adenine-competent state. At saturating adenine concentrations, the formation of this state leads rapidly to an adenine-bound, folded state that is long-lived even at ~7 pN (Fig. 4A), and the P1-unfolded state is thus rarely occupied. In contrast, absent adenine, the P1-unfolded state is frequently occupied even at low forces. Occasionally, the transient folding of P1 was observed even with adenine present (Fig. 4A, green), likely indicating that adenine was not bound at that instant. The single-molecule records thus directly reflect an adenine-induced stabilization of helix P1 that underpins the switching action of the riboswitch (28).

Each of the folding transitions can be analyzed individually as a two-state process, enabling a piecewise reconstruction of the energy landscape for folding, both with and without adenine (Fig. 4E). The relative free energies of the five observed states were determined from extension histograms, and the locations and heights of the energy barriers between states were determined from the force dependence of the kinetics (15, 18). From these landscapes, we find that the tertiary contacts that form the adenine-competent state, which are primarily base-pair and base-quartet interactions between hairpin loops (4, 9), stabilize the structure by an additional 2.7 ± 0.3 kcal/mol (18). The transition state for breaking these interactions lies ~1 nm from the adenine-competent state, indicative of their short range (29). We also find that adenine binding stabilizes the folded state by 4 ± 1 kcal/mol and raises the energy barrier for leaving the folded state, but does not appreciably affect other properties of the landscape.

These energy landscapes dramatically illustrate the sequential folding of each structural element in the RNA. Folding proceeds through a distinct hierarchy of states, but the formation of tertiary and secondary structure is interleaved, because the energetic stabilities of these structures happen to be comparable, in contrast with the standard picture of hierarchical folding. Indeed, the tertiary contacts that preorganize the adenine-competent state are considerably more stable than the least-stable helix, P1, which is the essential component governing the switching behavior of the riboswitch. In vivo, without adenine binding to stabilize P1, this last component of the aptamer to fold would be highly susceptible to disruption by terminator hairpin invasion.

The techniques developed here point the way to a powerful method for monitoring cotranscriptional folding. In the case of the *pbuE* aptamer, the first FEC obtained after transcription did not exhibit an unfolding behavior substantially different from that of subsequent FECs, implying that the cotranscriptional aspect of folding may not be important for the formation of an isolated aptamer (18). This result is unsurprising, because structural elements of the aptamer fold in the same order as they are transcribed; hence, force-induced refolding mimics cotranscriptional folding in this case. However, for the folding of the complete riboswitch, which includes a downstream terminator hairpin that competes with aptamer formation, we anticipate an important cotranscriptional dependence (10, 26).

References and Notes

1. W. C. Winkler, R. R. Breaker, *Annu. Rev. Microbiol.* **59**, 487 (2005).
2. R. L. Coppins, K. B. Hall, E. A. Groisman, *Curr. Opin. Microbiol.* **10**, 176 (2007).
3. R. T. Batey, S. D. Gilbert, R. K. Montange, *Nature* **432**, 411 (2004).
4. A. Serganov et al., *Chem. Biol.* **11**, 1729 (2004).
5. M. Mandal et al., *Cell* **113**, 577 (2003).
6. J. Noeske et al., *Proc. Natl. Acad. Sci. U.S.A.* **102**, 1372 (2005).
7. S. D. Gilbert, C. D. Stoddard, S. J. Wise, R. T. Batey, *J. Mol. Biol.* **359**, 754 (2006).
8. J. Noeske et al., *Nucleic Acids Res.* **35**, 572 (2007).
9. O. M. Ottink et al., *RNA* **13**, 2202 (2007).
10. J. F. Lemay et al., *Chem. Biol.* **13**, 857 (2006).
11. M. Mandal, R. R. Breaker, *Nat. Struct. Mol. Biol.* **11**, 29 (2004).
12. J. K. Wickiser, M. T. Cheah, R. R. Breaker, D. M. Crothers, *Biochemistry* **44**, 13404 (2005).
13. E. A. Abbondanzieri, W. J. Greenleaf, J. W. Shaevitz, R. Landick, S. M. Block, *Nature* **438**, 460 (2005).
14. I. Tinoco Jr., C. Bustamante, *Biophys. Chem.* **101-102**, 513 (2002).
15. M. T. Woodside et al., *Science* **314**, 1001 (2006).
16. M. T. Woodside et al., *Proc. Natl. Acad. Sci. U.S.A.* **103**, 6190 (2006).
17. R. V. Dalal et al., *Mol. Cell* **23**, 231 (2006).
18. Materials and methods are available as supporting material on *Science* Online.
19. B. Onoa et al., *Science* **299**, 1892 (2003).
20. O. K. Dudko, G. Hummer, A. Szabo, *Phys. Rev. Lett.* **96**, 108101 (2006).
21. S. B. Smith, Y. Cui, C. Bustamante, *Science* **271**, 795 (1996).
22. Y. Seol, G. M. Skinner, K. Visscher, *Phys. Rev. Lett.* **93**, 118102 (2004).
23. W. Saenger, *Principles of Nucleic Acid Structure* (Springer, New York, 1984).
24. J. Liphardt, S. Dumont, S. B. Smith, I. Tinoco Jr., C. Bustamante, *Science* **296**, 1832 (2002).
25. C. Jarzynski, *Phys. Rev. Lett.* **78**, 2690 (1997).
26. J. K. Wickiser, W. C. Winkler, R. R. Breaker, D. M. Crothers, *Mol. Cell* **18**, 49 (2005).
27. W. J. Greenleaf et al., *Phys. Rev. Lett.* **95**, 208102 (2005).
28. Some molecular heterogeneity was present, but we did not observe order-of-magnitude variations in the kinetics from one RNA to the next, as reported in one single-molecule fluorescence study (10).
29. P. T. X. Li, C. Bustamante, I. Tinoco Jr., *Proc. Natl. Acad. Sci. U.S.A.* **103**, 15847 (2006).
30. We thank R. Landick for providing purified RNAP, K. Herbert for providing the plasmid used to make the transcription template, M. Larson and O. Dudko for helpful advice, and members of the Block lab and Program Project grant GM066275 for useful discussions. Our research was supported by grants from the National Institute of General Medical Sciences (GM057035) and the National Institute for Nanotechnology.

New Products

Automated Microarray Processing

The QuadChamber is for automated processing of four different microarrays simultaneously on one slide using the HS Pro automated hybridization station. The QuadChamber was developed for use with Agilent's new 4 x 44 4-Plex Gene Expression as well as CGH Microarrays, which consist of four individual, whole-genome microarrays printed on a single glass slide. It is the first fully automated system that can independently handle four arrays on one slide with no cross-contamination between the arrays. The QuadChamber provides a sealed environment around each of the four arrays on Agilent's 4-Plex slides, with independent channels for wash buffers, independent agitation mechanisms, and independent drying.

Tecan Group
For information +41 44 922 81 11
www.tecan.com

HILIC Chromatography Columns

A new line of Luna HILIC (hydrophilic interaction liquid chromatography) columns for the analysis of polar metabolites in drug discovery and development delivers superior retention of polar compounds, increases mass spectrometry sensitivity, and increases laboratory throughput. The new columns retain a water-enriched layer on the surface of the silica that facilitates the transfer of polar compounds into the stationary phase for increased retention. Low-level polar metabolites and impurities have historically slowed development and negatively affected mass spectrometry sensitivity, and the new columns offer advancements in the separation, quantitation, and qualification of difficult-to-retain polar compounds.

Phenomenex
For information 310-212-0555
www.phenomenex.com

Antibody-Free Immunoprecipitation

The HaloCHIP System is the first antibody free alternative to chromatin immunoprecipitation (ChIP). ChIP is used to determine where, in vivo, a given protein binds to DNA. The major challenge of the ChIP method is the requirement for highly specific antibodies for each protein to be tested. In addition to working without the use of antibodies, the HaloCHIP system delivers results in less time than a standard ChIP (two days vs. four days). The new system also improves signal-to-noise ratios, enabling higher sensitivity for detection of even small changes in protein binding patterns from a reduced number of cells. The system is based on the HaloTag reporter protein, which has been engineered to form a covalent interaction with a series of specific ligands, including one for immobilization, HaloLink Resin. In HaloCHIP, DNA-binding proteins of interest are expressed in cells as HaloTag fusion proteins, cross-linked to DNA with formaldehyde, and then covalently captured on HaloLink Resin. After the captured complexes are washed to remove any nonspecific interactions, cross-links are reversed, and the specific DNA is released for downstream analysis.

Promega
For information 608-274-4330
www.promega.com

Spectrometer Family

The DNS spectrometers offer performance and flexibility at an affordable price. The series runs from the compact, short-focal-length DNS-150 with 0.4 nm resolution through the DNS-750, with a 0.03 nm resolution. All DNS systems support interchangeable grating turrets and are designed to provide high levels of mechanical stability and repeatability. It is easy to configure the DNS spectrometer with popular charge-coupled device cameras, including the MOSIR series; single-point detectors are also supported.

Intevac
For information 408-588-2150
www.intevac.com

Automated Cell Culture

The Vi-CELL XR Cell Viability Analyzer has been integrated with Biomek laboratory automation workstations and other devices for real-time cellular imaging in a walk-away system for cell culture settings. The Vi-CELL XR measures the viability of cells and counts them in minutes, using the widely accepted trypan blue cell viability tissue culture protocol. The instrument measures 15 to 30 times more volume than competitive systems in the same amount of time, with a more comprehensive array of parameters, according to the manufacturer. The automated, integrated system monitors cell growth over time.

Beckman Coulter
For information 714-993-8955
www.beckmancoulter.com

Plant DNA Extractions

QuickExtract Plant DNA Extraction Solution can be used to obtain genomic DNA ready for polymerase chain reaction (PCR) from most plant samples using a simple eight-minute protocol with two sequential heating steps. Most leafy plants are amenable to DNA extraction using the system. The method allows for the processing of samples without centrifugation, spin columns, or any toxic organic solvent. The resulting extract can be used for standard endpoint or real-time PCR applications.

Epicentre Biotechnologies
For information 800-284-8474
www.EpiBio.com

Electronically submit your new product description or product literature information! Go to **www.sciencemag.org/products/newproducts.dtl** for more information.